Digital Footprints as Collateral for Debt Collection

Lili Dai, Jianlei Han, Jing Shi, and Bohui Zhang[☆]

Current Version: August 2020

[☆] Lili Dai is from UNSW Business School, UNSW Sydney, Sydney, Australia; Jianlei Han and Jing Shi are from Macquarie Business School, Macquarie University, Sydney, Australia; Bohui Zhang is from School of Management and Economics and Shenzhen Finance Institute, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), Longxiang Boulevard, Longgang District, Shenzhen, China, 518172. Authors' contact information: Lili Dai, lili.dai@unsw.edu.au; Jianlei Han, jianlei.han@mq.edu.au; Jing Shi, jing.shi@mq.edu.au; Bohui Zhang, bohuizhang@cuhk.edu.cn. We are grateful for the valuable comments from Yi-Chun Chen, Junshan Duan, Yan Guo, Guangzhong Li, Yongxiang Wang, and the conference and seminar participants at Guangdong University of Finance and Economics, Macquarie University, and the 2020 Summer Research Workshop of China Journal of Accounting Research.

Digital Footprints as Collateral for Debt Collection

Abstract

We study the role of borrowers' digital footprints in debt collection. Using a large sample of personal loans from a fintech lender, we find that the lender can use the information in borrowers' digital footprints as collateral to facilitate debt collection. The effect is operated through bonding borrowers' obligations with their social networks and locating borrowers' physical locations. Moreover, the lender is more likely to approve loan applications from borrowers with digital footprints, although some of these borrowers can have a higher likelihood of delinquency. Our findings suggest that digital footprints, as a new type of collateral, can facilitate the collection of delinquent loans.

JEL classification: D14, G14, G23, G51

Keywords: Delinquency, Collateral, Debt collection, Digital footprints, FinTech

1. Introduction

Delinquency is a pervasive issue in consumer credit markets worldwide. According to the 2020 report of the Consumer Financial Protection Bureau (CFPB), outstanding credit card debt reached $927 billion in the U.S., and 8.4% was more than 90 days delinquent. Similarly, Chinese credit card holders accumulated a comparable amount of $13 billion overdue for more than six months by 2019. This enduring problem of loan delinquency calls for new tools to help creditors to obtain repayments. In recent years, the wide use of smartphones and internet services generates a new type of information about the online traces of users, such as the contacts that they maintain, the mobile applications that they use, and the websites that they visit. We refer to such information as digital footprints and propose the use of debtors' digital footprints potentially as a novel channel to facilitate creditors' debt collection.

Creditors' efforts in the loan process mainly include loan approval and debt collection. Recent research suggests that digital footprints are useful in evaluating borrowers' creditworthiness and thereby can assist creditors' loan approval decisions (Jiang et al., 2019; Agarwal et al., 2020; Berg et al., 2020).[1] Different from these studies, we consider the role of digital footprints in debt enforcement when loans are either delinquent or in default. In particular, we argue that the information acquired by lenders through borrowers' digital footprints can be used as collateral to recover the delinquent debt.

Pledged collateral has long been regarded by the debt literature as a commitment device to enforce repayment (e.g., Jimenez et al., 2006; Benmelech and Bergman, 2009; Rampini and Viswanathan, 2010, 2013; DeMarzo, 2019). However, as featured with unsecured loan contracting without any collateral, consumer credit often lacks such a guarantee. Especially in fast and small consumer credit contracts like those in marketplace lending, borrowers cannot provide any collateral assets. Thus, lenders need to look for other commitment technologies to secure debt claims (Thakor, 2020). In such scenarios, we view borrowers' digital footprints as a new form of collateral – *digital collateral*.

Digital collateral could be effective in the process of debt collection for two reasons. First, by using the information of borrowers' digital footprints, lenders may know exactly whom borrowers often contact via mobile phones, email, or instant messaging applications, and thus might be able to identify borrowers' social networks. When a loan is delinquent, the lender can

[1] Agarwal et al. (2020) and Berg et al. (2020) refer digital footprints to the information that borrowers leave online or through mobile devices when they access and register websites and mobile applications. Other studies on debt markets have examined the general role of technology-based lenders in mortgage lending markets (Fuster et al. 2019), the online shadow banks (Buchak et al., 2018), and the peer-to-peer (P2P) lending platforms (Tang 2019; Vallée and Zeng, 2019; Du et al. 2020).

reach the borrower's key contacts, such as family members or friends. This possibility may impair the borrower's trustworthiness among her social network, and more generally, her *social capital*. If this potential loss of the borrower's social capital is substantial, the borrower would be incentivized to repay the loan (e.g., Karlan et al., 2009).[2] Therefore, digital footprints as collateral could be used to bond borrowers' social capital with debt claims.

Second, through the use of online services, borrowers may also leave their address information in digital footprints, enabling lenders to identify their physical locations. In case of facing delinquency, lenders can use such information to facilitate debt collection by exercising liens to seize goods from borrowers' dwellings or other addresses. Specifically, for borrowers with no collateral in loan contracting, knowing their locations would be critical to enforcing loan repayments when lenders pursue judicial enforcement, and subsequently, judgment liens are placed on borrowers' assets by the courts.[3] Therefore, when delinquent borrowers receive chasing calls from debt collectors with the information of their addresses, they will perceive that the lenders will have the accessibility to their physical assets for debt enforcement and thus face pressures, motivating them to make loan repayments.

We examine our conjecture using the data of personal loan contracts from a leading fintech company in China. The fintech company only asks the individual borrower to provide her mobile number and national identity number. With such information, the lender can obtain from the third-party data providers other information about the borrower, such as the borrower's credit score, the balance of her online payment account, and the historical record of her past loan applications. Then, the lender computes the credit score using an internal automatic algorithm and evaluate the borrower's creditworthiness for loan decisions.

More importantly, this fintech creditor can also obtain the borrower's digital footprints from data providers and use them as collateral, including her frequent mobile contacts and online shopping delivery addresses. When the loan is delinquent, the fintech lender will make phone calls to both the borrower and her key contacts.[4] According to our hypothesis, the

[2] Karlan et al. (2009) develops a theory by adopting sociologists' concept of social capital, which "constitutes a particular kind of resource available to an actor" (Coleman, 1988), and "refers to features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit" (Putnam, 1995). Using the data on informal lending from Peru, Karlan et al. (2009) empirically documents the importance of bonding individuals' social capital to facilitate personal borrowings through social networks.

[3] Here, we refer the judgment liens to involuntary or non-consensual liens arise by statute or operation of the common law. When creditors seek legal actions for loan delinquency, liens will be imposed on borrowers' properties and chattels.

[4] Prior to the due date, the fintech lender sends reminder messages to borrowers via mobile phones (Cadena and Schoar, 2012; Karlan et al., 2015; Medina, 2018; Bursztyn et a., 2019; Du et al., 2020). In three months after the delinquency, if the borrowers still fail to make repayments, the lender will hand over the cases to debt collection agencies to further contact the borrowers and collect the debt (Fedaseyeu and Hunt, 2018; Fedaseyeu, 2020).

borrower's digital footprints would allow the lender to enforce debt collection by reaching the borrower's social network and/or locating her physical address. We thus expect borrowers with such digital collateral available to be more likely to repay their loans upon receiving the lender's phone calls.

Our empirical analyses start with a sample of 41,711 delinquent personal loan contracts over the period from July 2017 to November 2019. To examine the effect of digital collateral on the repayment of delinquent loans, we employ a difference-in-difference (DID) approach. Specifically, we first perform the propensity score matching (PSM) strategy by matching borrowers with digital collateral as the treatment group to those without collateral as the control group. This step provides us with two samples. One sample consists of 14,768 delinquent borrowers having frequent mobile contacts in their social networks, together with the same number of borrowers with no frequent contact information. Another sample includes 9,994 delinquent borrowers with online shopping delivery addresses and the matched ones without the address information.[5]

Second, in a timeline after delinquency, the lender will make the five rounds of phone calls to borrowers in the following quarter: the fourth day after the repayment due date (hereafter, Day 4), following by another four dates (i.e., Day 16, Day 31, Day 61, and Day 91). Then, the delinquent loans will be handed over to external debt collection agencies for further collection efforts after Day 91.[6] We examine the likelihood of repayment in several six-day windows centered around the dates when delinquent borrowers receive phone calls from the lender. This enables us to compare the changes in the likelihood of repayment after calls between the treatment and control groups.

Consistent with our conjecture, we find that after receiving the lender's phone calls, borrowers with digital collateral are more likely to repay their delinquent loans than those with no digital collateral. The effect is economically significant. For instance, in the first round of phone calls on Day 4, a delinquent borrower with frequent mobile contact (physical address) information has a 1.76% (1.79%) higher daily likelihood of loan repayment than those without such information from Days 4 to 6, compared to that of from Days 1 to 3. The economic magnitude accounts for an increase of 16.60% (18.54%) relative to the sample mean of the

[5] Frequent mobile contacts are defined as a borrower's mobile contacts who have made phone calls to the borrower for more than *ten* times, with each conversation longer than 20 seconds. For the online shopping delivery addresses, we exclude non-residential addresses, such as the delivery addresses of convenience stores in the local residential communities that provide the services of dropping and picking up the delivered shopping packages.

[6] See Figure 2 regarding the timeline. The fintech lender will send reminding messages from Days -3 to 0 (delinquency date), and then make chasing calls on Day 4, Day 16, Day 31, Day 61 and Day 91. After Day 91, the delinquent loans will be handed over to external debt collection agencies.

daily repayment likelihood in the six-day window centered on Day 4. The results remain qualitatively similar for the second round of chasing calls on Day 16.

Next, we conduct two placebo tests to validate our baseline findings. First, we re-perform the DID tests in a placebo sample period for the six-day window centered on the repayment due date, during which the lender makes no call but only sends reminders through mobile messages in three days before the due of repayment. We find no significant difference in the loan repayment likelihood between borrowers backed by digital collateral and those with no collateral. This finding indicates that prior to Day 4, when the first round of chasing calls is made, there is no difference between the treatment and control groups for a time trend of repayment around the due date. Second, we focus on the borrowers in control groups. Half of these borrowers are randomly assigned as the pseudo borrowers with digital collateral in new treatment groups, and the remaining half of them stay in new control groups, based on which we reconduct the DID analyses around Day 4. We again find insignificant results from this placebo test. Taken together, the findings from these placebo tests strengthen the inferences from our main analyses based on the DID empirical strategy for the effect of digital collateral on debt collection.

Moreover, we perform analyses based on the amount of digital collateral, further chasing calls, and alternative sample selections. We first decompose the metrics of digital collateral based on the median numbers of borrowers' frequent contacts and addresses. The results suggest a marginal decline in the coefficients on collateral proxies with a large amount of collateral compared to those with a small amount of collateral. This is likely because multiple pieces of digital footprints may distract debt collectors' efforts to borrowers' less important personal information, reducing the effectiveness of using digital collateral. Then, we examine the chasing calls in the next three rounds (i.e., Day 31, Day 61, and Day 91). We find that the effect of these further chasing calls on debt collection remains significant, but the magnitude of such an effect is significantly reduced. These results indicate that the borrowers, who continue to be delinquent after the first two chasing calls, are less concerned about their social capital and physical assets, suggesting an impairment of their digital collateral. Lastly, we examine whether our baseline findings are robust to alternative sample selections based on the overlap between the two types of digital footprints, i.e., key contact and physical location. We find our inferences remain unchanged for borrowers with only one type of digital footprint or those with both digital footprints.

Finally, to provide a complete picture of the loan process, we examine the effect of digital collateral on the repayment amount, delinquency likelihood, and loan approval. First, we study

the impact of the digital collateral on the magnitude of debt recovery after delinquency. Among the 41,711 delinquent loans, we find that the amount of debt repayment scaled by total debt outstanding is positively associated with the availability of borrowers' digital collateral. Furthermore, we focus on the dummy of debt recovery, indicating whether the repayment amount is equal to or more than total debt outstanding caused by the penalty of delinquency. We observe that the likelihood of debt recovery also increases with the availability of digital collateral. These results provide further confirmation on the role of digital footprints as collateral in debt collection.

Second, we examine whether the likelihood of loan delinquency is associated with the presence of borrowers' digital footprints. One possible concern on our headline findings is that the existence of frequent mobile contacts and physical addresses might proxy for borrowers' creditworthiness rather than collateral. For example, the availability of digital footprints would imply a low likelihood of delinquency and a high likelihood of repayment. To test this possibility, we investigate our sample of 97,783 approved loan contracts, consisting of both delinquent and non-delinquent loans. We show that the incidence of delinquency is even positively associated with the presence of borrowers' frequent contact information, while other conventional proxies for creditworthiness are negatively associated with the loan delinquency rate, consistent with prior studies (e.g., Berg et al., 2020). This finding helps to alleviate the concern that digital footprints are merely a measure of borrowers' credit quality.

Finally, we investigate the impact of digital collateral on the fintech lender's loan approval decisions based on a sample of 236,967 loan applications, including unapproved contracts. Intuitively, we find that the lender is more likely to approve the loan for a borrower with digital collateral available, after controlling for credit quality proxies such as the borrower's online credit score, the balance in her payment account, and the historical record of her past loan applications. This finding suggests an incremental effect of digital collateral on the top of the information regarding the borrower's credit quality. That is, the lender does take the borrower's digital footprint information into account when determining the loan approval, partly because the lender would be able to collect repayment effectively from the borrower in the case of delinquency.

Our findings make two contributions to the literature. First, our findings improve our understanding of the implication of digital footprints by documenting its collateral role in facilitating debt collection. Prior studies focus on the risk relevance of digital footprints for the credit assessment of borrowers, complementing the credit bureau scores, in debt markets. For example, Berg et al. (2020) shows that the borrowers using mobile devices from Apple have a

lower likelihood of loan default than those using devices based on Android operating systems. Agarwal et al. (2020) documents that digital footprint information can predict both the default rate and loan approval. Our paper contributes to the literature by showing that the information in digital footprints can be used as collateral to facilitate debt collection from delinquent debtors, from whom a marketplace lender typically cannot demand asset-backed security in personal loan contracts (Thakor, 2020). Furthermore, we shed light more broadly on the usage of big data analytics, which has been documented to improve price informativeness in the equity market and, consequently, discipline managers' behavior (Zhu, 2019; Grennan and Michaely, 2020).

Second, our paper adds to the literature on debt collateral and collection. Prior theoretical studies suggest the roles of collateral in mitigating the financial frictions (Johnson and Stulz, 1985; Aghion and Bolton, 1992; Hart and Moore, 1994, 1998), and incentivizing repayments in debt markets (Kiyotaki and Moore, 1997; Rampini and Viswanathan, 2010, 2013). Empirically, Benmelech and Bergman (2009) has documented that debt collateral can lower the credit spread, and Schiantarelli et al. (2020) finds that the debt collateral backed by assets can mitigate the likelihood of delinquency. Our study complements prior research by documenting the digital collateral backed by borrowers' social and physical capital, in enforcing debt repayments after delinquency.

The rest of the paper is organized as follows. Section 2 describes the sample selection, variable construction, and summary statistics of the variables used in our study. Sections 3 presents the main results for the association between digital footprint and debt collection. Section 4 reports the tests for other outcome variables of the loan process, and Section 5 provides further discussion. Section 6 concludes the paper.

2. Institutional background, variables, and descriptive statistics

In this section, we describe the institutional background of our research setting and present the definitions and summary statistics of variables in our main analyses.

2.1. Institutional background

We obtain the data of personal loan contracts from a leading fintech company in China with the lending business starting from 2017. To allow the potential borrowers to submit their loan applications, the company set up application terminals in retail stores across most major cities in China, with the staff assisting the loan applicant in the application process. Figure 1 illustrates the process of the loan application, approval, repayment, and collection.

[Insert Figure 1 Here]

A borrower can apply for the loan from the fintech company's terminals. The information filled in the loan application includes the customer's mobile number and national identity card. Then, by using the applicant's mobile number and national identity number, the company can obtain the applicant's personal data from third-party data providers. The third-party information covers the applicant's mobile phone call logs from mobile carriers, online shopping delivery addresses, credit score from Tencent (the largest social media company in China), Alipay account balance from Alibaba (the largest e-commerce company in China), and historical record of loan applications from other online lending companies.[7] With the information from the loan applicant's digital footprints and national identity card (e.g., age and gender), the fintech lender will determine whether the loan application should be approved. Then, the lender will inform the applicant with the decision of either loan approval or rejection within ten minutes.

After the loan application is approved, the borrower will obtain money from the fintech company, and then use it to purchase durable goods such as mobile phone immediately after the loan approval, or keep the cash for future use. Then, in the next six or twelve months, subject to the loan term's length, the borrower needs to pay the monthly principal and interest back to the lender. When a borrower fails to make a monthly repayment before the due date, the fintech lender will execute a standardized debt collection process against the delinquent borrower. Figure 2 illustrates the timeline of the company's debt collecting process.

[Insert Figure 2 Here]

As shown in the timeline, in a four-day window [-3, 0] before the repayment due date (Day 0), the lender starts to send mobile messages as reminders to borrowers. When a borrower does not make the repayment, on Day 4, the lender will call the borrower and tell her that debt delinquency may negatively impact her nationwide credit score, which is widely shared and used by financial institutions and public services in China. It would be hard for a person with a low credit score to obtain future mortgage loans and even travel with high-speed trains and airlines.

The lender will also make phone calls to the borrower's frequent mobile contacts if such information can be obtained from her digital footprints. When the lender has the borrower's

[7] The applicant's digital footprints can be directly sourced from online and mobile service providers or extracted by the data providers using crawler technologies from internet websites.

physical location in digital footprints, the lender will tell the borrower that debt collectors would come to the borrower's place to collect the loan and report it to the borrower neighborhood committee.[8] Further, the lender will warn the borrower with the penalties forced by the police and the potential lawsuit in court. When the borrower's address information is available, the lender will link the law enforcement to the borrower's physical location in the chasing call.

After Day 4, if the borrower still does not make the loan repayment in the coming quarter, the lender will call the borrower and her frequent contacts if available, again on Day 16, Day 31, Day 61, and Day 91. Finally, after the last chasing call (Day 91), the case of the delinquent loan will ultimately be handed over from the fintech lender to a third-party debt collection agency, who will make further efforts on loan collection.

2.2. Variable construction

2.2.1. Loan repayment and collection variables

Our sample starts with 236,967 personal loan applications from July 2017, when the fintech company started the lending business, to November 2019. Among all these loan applications, 97,783 applications have been approved. Figure 3 presents the geographical distribution of the loan application density, i.e., the number of applications per million of the population across the provinces in China. As indicated in the figure, the sample of applications is widely spread over the country and well represents the population in China.

[Insert Figure 3 Here]

Within 97,783 approved loans, 41,711 borrowers experience at least one time of loan delinquency. Our main analyses focus on first-time delinquency for these borrowers. To measure the outcome of loan repayment on a daily basis, we construct an indicator variable, $Paid_t$, equal to one if a borrower makes the repayment on the Day t, and zero otherwise, with Day 0 referred to the due date. We define $Chasing_{Day\,t}$ as an indicator variable equal to one for days in a three-day window $[t, t+2]$ after the lender calls delinquent borrowers and zero otherwise for the DID analyses. Day t refers to Day 4, Day 16, Day 31, Day 61, and Day 91. In our analyses for the association between chasing call and loan repayment, for a chasing call on Day t, we restrict the sample by only including borrowers who remain delinquent until Day

[8] In China, a neighborhood committee administers the dwellers living in the residential community. Reporting to the committee may enable the spread of a borrower's loan delinquency news, thereby damaging the borrower's social capital in the residential community.

t-3. For example, we study the borrowers who remain delinquent on Day 1 (Day 13) in the analyses for chasing calls on Day 4 (Day 16).

In additional analyses, we further construct the loan outcome variables as follows. $Repayment_{Ratio}$ is defined as the ratio of the repayment amount to total debt outstanding. $Repayment_{Complete}$ is an indicator variable for loan recovery if the repayment amount is equal to total debt outstanding, or more than the outstanding amount caused by the penalty of delinquency. *Delinquency* is an indicator variable that equals one if a loan is delinquent, and *Approval* is an indicator variable that equals one if a loan application is approved.

2.2.2. Digital collateral variables

We construct the two measures as the metrics of digital collateral in our main analyses. First, we define $Contact_{Existence}$ as the indicator variable for the existence of frequent mobile contacts. We apply a few filers to identify frequent contacts: 1) a frequent contact number should not be a phone number from commercial service; 2) a frequent contact should have made more than ten phone calls with a borrower; and 3) each phone call to be counted toward a frequent contact should have been made for more than 20 seconds.

Second, we construct the digital collateral metrics, $Address_{Existence}$, for the existence of physical address. The fintech lender acquires the address data through the borrower's online shopping transactions in Taobao, an online shopping site headquartered in Hangzhou City and owned by Alibaba. We filter out the addresses that neither seem to be the borrower's residential or office addresses nor are useful to identify the borrower's location. First, we exclude the addresses with keywords in relation to convenience stores that provide the services for customers to pick up the delivered shopping packages. Second, we further exclude the addresses with no street or unit number.

In further analyses, we define more digital collateral proxies in the following ways. We first construct two pseudo collateral measures, $Contact_{Pseudo}$ and $Address_{Pseudo}$, by randomly assigning borrowers with no collateral to treatment groups as if they have corresponding digital collateral. Then, we estimate the number of borrowers' frequent contacts, $Contact_{Number}$, and the number of borrowers' physical addresses, $Address_{Number}$. When a borrower has no collateral, we code $Contact_{Number}$ and $Address_{Number}$ as *zero*. Finally, we decompose $Contact_{Number}$ ($Address_{Number}$) by the median value of contact number (address number) in our sample, into $Contact_{1 \leqslant Number \leqslant 7}$ and $Contact_{Number>7}$ ($Address_{Number=1}$ and $Address_{Number>1}$). For example,

Contact $_{1 \leqslant Number \leqslant 7}$ indicates whether a borrower has at least one and at most seven frequent contacts.

2.2.3. Control variables

We have three sets of control variables in our regression analyses. First, we include loan characteristics in the model. Berg et al. (2020) finds that the loan amount is negatively associated with the default likelihood, potentially due to the notion that borrowers who can obtain a large amount of loan tend to have better credit scores. We control the amount of loan principal in thousands of Chinese Yuan (*Amount*), which is transferred to the logarithm value in regressions. The prediction of this control variable is ambiguous. On the one hand, it is possible that borrowers with larger loan amounts are more likely to make repayments, because of higher creditworthiness suggested by Berg et al. (2020). On the other hand, these borrowers in our debt collection analyses have experienced loan delinquency; therefore those with high credit quality may have already been filtered out in the first place. We also control for the interest rate of loan on an annual basis (*Rate*), as Iyer et al. (2016) has documented that interest rates of loans are correlated with borrowers' credit scores, both predicting loan performance in terms of default and repayment.

Second, we control for the borrower's personal characteristics, including the age of a borrower in years (*Age*), which is transferred to logarithm value is taken in regression analyses, and the indicator variable for the female borrower (*Gender*). Lusardi and Mitchell (2014) documents the gender differences in financial literacy around the world, which could lead to different repayment behaviors between female and male borrowers in our analyses. Berg et al. (2020) shows that older borrowers are less likely to default, suggesting a credit history built up over time, but finds an insignificant gender effect on the default likelihood.

Third, we directly control for the borrower's creditworthiness metrics (Iyer et al., 2016; Berg et al., 2020), such as the credit risk score provided by *Tencent*, with a high value suggesting high-risk profile (*Score*), the balance of a borrower's *Alipay* account from *Alibaba* in Chinese Yuan (*Wealth*), and the number of loan applications of a borrower rejected by other online lending platforms (*History*). All these credit quality proxies are transferred to logarithm values in our regressions. We have no strong prior belief about whether these credit measures would predict the loan repayments for delinquent borrowers since they have defaulted once on their loan payments.

2.3. Descriptive statistics

Table 1 presents the summary statistics of the main variables used in our analyses. We find that the average likelihood of loan delinquency, *Delinquency*, is 42.7% in a sample of 97,783 loans (41,711 / 97,783), and that of loan approval is 41.3% in a sample of 236,967 loan application (97,783 / of 236,967). For 41,711 delinquent borrowers, 61.9% of them make repayments ($Paid_{[1,6]}$ = 61.9) within a six-day window [1, 6] after the first round of chasing calls. The average likelihood increases to 75.2% and 84.7% in an 18-day window [1, 18] and a 93-day window [1, 93], respectively.

The average ratio of repayment to total debt outstanding, $Repayment_{Ratio}$, is 1.07, and the average recovery likelihood, $Repayment_{Complete}$, is 73.8%, lower than the repayment likelihood in the window [1, 93]. This suggests that not all repayments fully recover the outstanding loan payments. Indeed, in untabulated statistics, we find that among all the repayments made to delinquent loans, 12.0% (73.8%) borrowers make underpayments (overpayments), accounting for 29.0% (123.6%) of the outstanding loans. The remaining 14.2% repayments are made on par with the outstanding amounts.[9]

Turning to the digital collateral metrics, 65.1% delinquent borrowers have the information of frequent mobile contacts associated with the mobile numbers provided to the fintech lender ($Contact_{Existence}$ = 1), and the counterpart group includes 34.9% borrowers with no contact that makes more than ten phone calls ($Contact_{Existence}$ = 0), lending to a fairly large variation for our analyses on digital collateral. For a typical borrower, the mean value of $Contact_{Number}$ is 7.49, and the median and upper quartile values are 3.00 and 9.00, suggesting a right-tailed skewness distribution. With regard to the physical address, the lender can find the address information in digital footprints of 76.5% delinquent borrowers ($Address_{Existence}$ = 1). The mean and median values of $Address_{Number}$ are 1.06 and 1.00, respectively.

The average principal amount of personal loans is 3,966 Chinese Yuan, which is equivalent to about 567 U.S. dollars. This amount can be economically significant to an average borrower in China, considering the comment by the Premier of China, Keqiang Li, at the close of China's Two Sessions congressional meeting on 28 May 2020. That is, "600 million people have monthly incomes of just 1,000 Yuan" (around 143 U.S dollars), who are likely to borrow cash to complement their low incomes through fast and small lending contracts, typically with no collateral. The annual nominal interest rate is 29.5%, substantially higher than the prime

[9] There is a likelihood of overpayment because the fintech lender charges the delay-repayment penalties on the delinquent borrowers.

loan rate at 4.3% between 2017 to 2019, but lower than the cut-off rate for the usurious loan equal to 36.0% as per statute in China.[10]

The average age of borrowers is 27.1, with the top quartile equal to 31 and the bottom quartile equal to 21, representing a group of young borrowers who are generally familiar with the use of online services, and likely to leave their digital footprints through the use of electronic devices. Only about 18.2% of delinquent borrowers are female. Compared to the untabulated female percentage in the full sample equal to 19.2%, this suggests that male borrowers are more likely to experience loan delinquency. On average, a typical borrower has the credit score at Tencent equal to 56.8, personal wealth in Alipay's account amounting to 4.22 Chinese Yuan, and has been rejected by other online lending platforms by 0.12 times.

[Insert Table 1 Here]

3. Main results

Our main analysis is conducted using a difference-in-difference approach by matching borrowers with digital collateral available as the treatment group to those with no collateral as the control group. We conduct the tests based on the two matched samples, one for borrowers with or without the frequent contact information, and the other for borrowers with or without the physical address information.

3.1. Matched samples

We employ the propensity score matching strategy to construct our matched samples on a one-to-one basis without replacement. The PSM approach accounts for both loan characteristics (*Amount* and *Rate*) and borrower characteristics (*Age*, *Gender*, *Score*, *Wealth*, and *History*), which may influence the debt collection process. Given that there are more delinquent borrowers with digital collateral than those with no collateral (e.g., the mean of $Contact_{Existence}$ = 65.1%), it is likely that a treatment borrower would have no match in control group when her distance of propensity score to any control borrower is not the least compared to all other

[10] There are two cut-off rates for the usurious loans in China, 24% and 36%. For example, when a borrower enters a loan contract with an annual interest rate equal to 30%, the interest proportion below 24% needs to be paid by the borrower, but the proportion between 24% and 30% will not be protected by the law. However, if the borrower has paid the interest between 24% and 30%, she cannot require the lender to return the interest payment. In another example, if a borrower enters a loan contract with an annual interest rate equal to 40% and has paid the interest, she has a right to ask the lender to return the interest proportion between 36% and 40%, and this request will be supported by the court.

treatment borrowers.[11] This approach provides us with two matched samples. One sample consists of 14,768 delinquent borrowers having frequent mobile contacts in their social networks as the treatment group, together with the same number of borrowers with no frequent contact information as the control group. Another sample includes 9,994 delinquent borrowers with online shopping delivery addresses as the treatment group and the ones without the address information as the control group.

Table 2 presents the comparison of loan and borrower characteristics between the treatment and control groups. In Panel A, we focus on a treatment group of 14,768 borrowers with frequent contact information and find no statistical difference in loan and borrower characteristics between the two groups. This finding validates that our PSM process is well executed. In Panel B, we find similar results for a treatment group of 9,994 borrowers with physical addresses, except that the treatment borrowers encounter higher interest rates (*Rate*) and have lower balances of online payment accounts (*Wealth*), compared with control borrowers. Although the significant differences between the two groups may suggest the imperfection of our matching process, these differences are biased against our main analyses. This is because we conjecture that the treatment borrowers with digital collateral are more likely to make the repayments, while the higher interest rates and lower wealth balances, on the contrary, can make it more difficult for borrowers to repay the outstanding loans.

[Insert Table 2 Here]

3.2. Univariate analyses

We conduct the univariate analyses in Table 3 along the timeline of loan delinquency. In Panel A, we find that the likelihood of repayment is 11.68% in a three-day window [1, 3] after the delinquency for the treatment borrowers with the information of frequent mobile contacts ($Contact_{Existence} = 1$) in a sample of 44,304 borrower-day observations for 14,768 delinquent borrowers. The likelihood is 11.78% for the control borrowers with no frequent contact information ($Contact_{Existence} = 0$). The difference between the two groups is insignificant.

After the fintech lender makes calls made on Day 4, the treatment borrowers have a significantly higher repayment likelihood (9.51%) in the window [4, 6] than the control borrowers (7.83%). This is because the lender tells treatment borrowers and their frequent

[11] We find similar results when we perform the PSM strategy on a one-to-one basis *with* replacement, resulting in matched samples in which each treatment borrower can be matched to a borrower in the control group, while a control borrower can be re-used matching to multiple borrowers in the treatment group.

contacts that debt delinquency may cause a negative impact on borrowers' nationwide credit scores, but only warns the control borrowers themselves. The difference in the likelihood between the two groups is equal to 1.68% and statistically significant at one percent level. Economically, moving from the control group to the treatment group, there will be an increase in the repayment rate for a magnitude of 21.46% (1.68% / 7.83%).

[Insert Table 3 Here]

Our findings are similar around the second wave of phone calls on Day 16 for 4,412 borrowers who remain delinquent on Day 13. Specifically, the lender will warn the treatment borrowers ($Contact_{Existence} = 0$) and their frequent contacts again, as opposed to the control group in which only borrowers will be called. The results show that in a three-day window [13, 15] prior to the phone call, there is no significant difference in the repayment likelihood between the treatment and control groups, however, in the post-call window [16,18], borrowers with the frequent contact information have a significantly higher repayment rate (3.68%) than those with no such information (2.52%), with the difference equal to 1.16% and significant at one percent level. This finding is also economically significant. Compared to borrowers with no frequent contact, the treatment borrowers have an increase in repayment rate for 46.03% (1.16% / 2.52%).

The further DID analyses suggest that the differences between the pre- and post-windows are also significant for the two groups of borrowers. For example, we find that although in the post-window [4, 6] there are reductions in the repayment rates for both the treatment and control groups, the decrease in the repayment likelihood is lower for the borrowers with the frequent contact information (−2.17%, $Contact_{Existence} = 1$) than the borrowers without such information (−3.95%, $Contact_{Existence} = 0$).

Moreover, these results also hold in Panel B for the 9,994 treatment borrowers with physical addresses for the first round of chasing calls on Day 4. The difference between borrowers with physical addresses ($Address_{Existence} = 1$) and those with no address ($Address_{Existence} = 0$) is -0.24% in the pre-call window [1, 3], while this difference is 1.57% in the post-call window [4, 6] and significantly positive at the one percent level. The difference in the pre-post difference between the two groups is 1.80%, also significant at the one percent level. This is because the lender will tell the treatment borrowers that the lender's staff would come to their places for debt collection, report the delinquent news to their neighborhood committees, and possibly take further legal actions to enforce loan repayments at the borrower's physical

address. However, the lender will only warn the control borrowers on the potential damages on their credit scores.

Lastly, we examine the second wave of calls, when the lender will call the treatment borrowers ($Address_{Existence} = 1$) with physical addresses again on Day 16, and find similar results for 3,459 borrowers who remain delinquent until Day 13. The patterns of prepayment likelihood discussed above have also been illustrated in Figure 4 and Figure 5. From both figures, we find the larger increases in periods after the fintech lender makes the chasing calls to digital-collateral-backed borrowers than those with no digital collateral.

[Insert Figure 4 and Figure 5 Here]

3.3. Main findings

In this sub-section, we examine the effect of digital collateral through chasing calls on debt collection using the multivariate regression analysis. We specify our main DID analysis within a six-day window [*s-3*, *s+2*] in the following model, with Day *s* indicating the date for the lender to make chasing calls:

$$Paid_t = \alpha + \beta_{CD\times DC}\, Chasing\ Day_s \times Digital\ Collateral\ Proxy + \beta_{CD\times Control}\, Chasing\ Day_s \times Control\ Variables + \beta_{FE}\, Fixed\ Effects + \varepsilon, \quad (1)$$

where $Paid_t$ indicates if the repayment is made by a borrower on Day *t*, and $Chasing\ Day_s$ is equal to one if Day *t* is in the three-day window [*s*, *s+2*], and zero if Day *t* is in the window [*s-3*, *s-1*]. *Digital Collateral Proxy* denotes the metrics of digital collateral constructed based on digital footprints, i.e., $Contact_{Existence}$ and $Address_{Existence}$, constructed for each borrower. We interact $Chasing\ Day_s$ with *Digital Collateral Proxy* to capture the DID effect of *Digital Collateral* between the pre- and post-chasing-call periods. We expect the interaction term's coefficient to be significantly positive when digital collateral indeed plays an effective role in facilitating debt collection.

We conduct the analyses in the matched samples, as described in Table 3. In Equation (1), *Control Variables* represents a vector of control variables of loan and borrower characteristics, as described in Section 2.2, which are interacted with $Chasing\ Day_s$ to account for any effect of chasing calls that moderate the impact of control variables on the repayment likelihood. We

also include *Fixed Effects*, the borrower and day fixed effects, to control for any cross-sectional and time-series omitted factors and cluster standard errors at the borrower level.[12]

We perform the DID tests based on Equation (1) for the first two rounds of chasing calls in Table 4. Around the first chasing call on Day 4, for the 14,768 treatment borrowers with frequent mobile contacts, the matched sample in the DID analyses consists of 177,216 borrower-day observations for both treatment and control groups in the six-day window (14,768 × 2 × 6). For the 9,994 treatment borrowers with the information of physical addresses, the matched sample being tested include 119,928 borrower-day observations (9,994 × 2 × 6). Similarly, we examine the second round of chasing calls on Day 16 for the remaining 4,412 and 3,459 treatment borrowers, who have not made repayment until Day 13, with contact and address information. The two samples for the second chasing call tests consist of 52,944 (4,412 × 2 × 6) and 41,508 borrower-day observations (3,459 × 2 × 6), respectively.

The results suggest the usefulness of digital collateral in the debt collection process. For example, in Model 1, we find that delinquent borrowers with frequent contact information have a significantly larger increase in the repayment likelihood, from the pre-call window [1, 3] to the post-call window [4, 6], than those with no frequent contact information. The coefficient on $Chasing_{Day\ 4} \times Contact_{Existence}$ is equal to 1.76, and significant at the one percent level, which accounts for an increase of 16.60% relative to the average daily repayment likelihood in the window [1, 6].

Regarding our control variables, we find a significantly positive impact of chasing calls on relatively old borrowers (coefficient on $Chasing_{Day\ 4} \times Age$ = 5.71 and *t*-statistic = 10.21) and borrowers with high credit risk profiles (coefficient on $Chasing_{Day\ 4} \times Score$ = 3.35 and *t*-statistic = 6.50). This potentially indicates that old and risky borrowers are more concerned about the possible detrimental impacts on their credit scores if they do not make the repayments. In contrast, there is a significant and negative reciprocal effect of chasing calls on loan repayment for loans with large amounts (coefficient on $Chasing_{Day\ 4} \times Amount$ = -1.92 and *t*-statistic = -4.62) and high interest rates (coefficient on $Chasing_{Day\ 4} \times Rate$ = -12.74 and *t*-statistic = -6.98), and also for female borrowers who could be more annoyed by chasing calls and therefore are less motivated to repay the delinquent loans (coefficient on $Chasing_{Day\ 4} \times Gender$ = -2.29 and *t*-statistic = -5.72).

[12] In this DID specification, we do not control for the main effect of Chasing Day *s*, because this effect has been absorbed by the day fixed effects. We do not control for the main effects of control variables, because the interaction terms have absorbed them. We find similar results when we drop the interactions and only control for the main effects of control variables.

We find similar results across when we examine $Address_{Existence}$ and investigate the effect of digital collateral for the second chasing call on Day 16. In Model 2, the coefficient on $Chasing_{Day\ 4} \times Address_{Existence}$ is 1.79 and significant at one percent level (t-statistic = 5.00). In Model 3 (Model 4), the coefficients on $Chasing_{Day\ 16} \times Contact_{Existence}$ and $Chasing_{Day\ 16} \times Address_{Existence}$ are 1.03 and 0.90 and both significant at one percent level (t-statistics = 3.67 and 3.04). However, these coefficients are magnetically smaller than those in Models 1 and 2 with z-statistics of differences in coefficients equal to 1.76 and 1.93, respectively, suggesting that borrowers remain delinquent after the first round of calls are less responsive to further calls, because they are less concerned about their social capital and being chased by debt collectors through their physical locations.

Taken together, the results in Table 4 provide supports for an effective role of digital collateral in debt collection.

[Insert Table 4 Here]

3.4. Placebo tests

One potential concern regarding our main findings is that the DID analyses may capture a time trend effect around the chasing calls, coinciding with the presence of digital collateral. In this sub-section, we run two placebo tests to alleviate this concern.

First, we re-run the DID analysis in a six-day window [-2, 3] by considering Day 1 as the placebo chasing call date. If our previously documented results for chasing calls made on Day 4 are driven by a time trend effect, we should observe a similar effect of digital collateral in this placebo test. That is, borrowers with the information about frequent mobile contact and physical address have an increase in the repayment likelihood from the window [-2, 0] to [1, 3] than those with no such information. In Table 5, we find that the coefficients on $Chasing\ Day_{1} \times Contact_{Existence}$ and $Chasing\ Day_{1} \times Address_{Existence}$ are -0.074 and -0.266 and insignificant (t-statistics = -0.40 and -1.22).

[Insert Table 5 Here]

Second, we run the placebo tests based on pseudo borrowers with digital collateral. We only focus on the borrowers in control groups of our main tests, who have no digital collateral ($Contact_{Existence}=0$, or $Address_{Existence}=0$). Then, we randomly assign half of the borrowers in each control group as the pseudo borrowers with digital collateral and define new indicator variables equal to one for these pseudo borrowers, $Contact_{pseudo}$ and $Address_{pseudo}$,

respectively. The unassigned borrowers remain in the control groups with $Contact_{pseudo}$ and $Address_{pseudo}$ denoted zero. We reconduct the DID analyses based on the pseudo borrowers and the chasing calls made on Day 4. The results are reported in Table 6. The sample size is reduced to 88,608 and 59,964, which is the half of the sample size in Table 4 (i.e., 177,216 and 119,928), because we essentially focus on borrowers with no collateral in the placebo tests. We find that the coefficients on $Chasing_{Day\ 4} \times Contact_{pseudo}$ and $Chasing_{Day\ 4} \times Address_{pseudo}$ are -0.467 and -0.155 and insignificant (*t*-statistics = -1.13 and -0.32).

[Insert Table 6 Here]

Collectively, these findings in these placebo tests suggest that the time trend effect cannot be the main force that predominantly drives our baseline findings, and therefore strengthen the inferences associated with the digital collateral effect in our main analyses.

3.5. Alternative collateral proxies, time windows, and sample selections

In this sub-section, we perform the analyses using alternative proxies of digital collateral, different time windows based on further chasing calls, and sub-samples for borrowers with or without overlap between the two types of digital footprints.

3.5.1. Analyses based on the amount of digital collateral

In Table 7, we replace our main metrics of digital footprints regarding the presence of borrowers' contact and location information with the amount of digital information, to investigate whether the variation of information amount exerts an effect on debt collection.

In Models 1 and 2. We employ the numbers of a borrower's frequent contacts and physical addresses (i.e., $Contact_{Number}$ and $Address_{Number}$) as the main independent variables of interest, and re-conduct our DID analyses. Consistent with our main results, we find a positive impact of $Contact_{Number}$ and $Address_{Number}$ on the likelihood of loan repayment through chasing calls. The coefficients on $Chasing_{Day\ 4} \times Contact_{Number}$ and $Chasing_{Day\ 4} \times Address_{Number}$ are 0.51 and 1.84 and both significant at one percent level (*t*-statistics = 3.93 and 4.43). These findings indicate an average positive effect of the digital footprint amount on loan repayment.

Also, as discussed in Section 2, we decompose $Contact_{Number}$ and $Address_{Number}$ into sub-components, i.e., $Contact_{1 \leqslant Number \leqslant 7}$, $Contact_{Number>7}$, $Address_{Number=1}$, and $Address_{Number>1}$, based on the sample medians. In Model 3, we find that the coefficients on $Chasing_{Day\ 4} \times Contact_{1 \leqslant Number \leqslant 7}$ and $Chasing_{Day\ 4} \times Contact_{Number>7}$ are 2.13 and 1.31 and both significant at one percent level (*t*-statistics = 5.96 and 3.30). The difference between the two coefficients

is significant at ten percent level (z-statistic = 1.83). One possible explanation is that for a borrower with many frequent contacts, the lender may choose some contacts to make chasing calls rather than reaching all the contacts because of time and resource constraints. However, these selected contacts can be less important than those for a borrower with a small number of key contacts. In Model 4, we find a large coefficient on $Chasing_{Day\ 4} \times Address_{Number=1}$ (coefficient = 1.85, t-statistic = 4.74) relative to the coefficient on $Chasing_{Day\ 4} \times Address_{Number>1}$ (coefficient = 1.65, t-statistic = 2.89), although the difference between the two coefficients is statistically insignificant.

In sum, the results suggest that the likelihood of loan repayment increases with the amount of digital collateral, and this increase may not be linear. Especially when there is a large number of borrowers' digital footprints, the lender can potentially spend more time identifying the key information, which, in turn, would reduce the effectiveness of the digital collateral use.

[Insert Table 7 Here]

3.5.2. Analyses based on further chasing calls

We conduct the analyses based on the next three rounds of chasing calls on Day 31, Day 61, and Day 91, as illustrated in Figure 2, to examine the impacts of using digital collateral along the whole timeline before the lender hands over delinquent loans to external collection agencies.

In particular, we perform the DID tests in five six-day windows for all the chasing calls. The number of treatment borrowers with the frequent contact information in digital footprints, who remain delinquent in three days before each chasing call (i.e., on Day 1, Day 13, Day 28, Day 58, and Day 88), is 14,768, 4,412, 3,346, 2,638, and 2,388 for calls on Day 4, Day 16, Day 31, Day 61, and Day 91, respectively. Therefore, the sample testing the frequent contact information for all the chasing calls includes 330,624 borrower-day observations.[13] For treatment borrowers with physical location information, the sample consists of 241,812 observations, because the number of borrowers with the frequent contact information varies across the five calling days equal to 9,994, 3,459, 2,696, 2,102, and 1,900, respectively.[14]

Table 8 presents the results based on all the chasing calls. We first construct an indicator variable, $Chasing_{All}$, equal to one for days in windows [4, 6], [14, 16], [31, 33], [61, 63], and [91, 93], and zero otherwise. Models 1 and 2, we find an overall effect of digital collateral,

[13] The sample for treatment borrowers with contact information consists of (14,768 + 4,412 + 3,346 + 2,638 + 2,388) × 2 × 6 = 330,624 borrower-day observations.
[14] Accordingly, the sample size is equal to (9,994 + 3,459 + 2,696 + 2,102 + 1,900) × 2 × 6 = 241,812.

facilitating the debt collection process across all the five six-day windows. For example, in Model 1, we find that in general, the borrowers with frequent contact information are more likely to repay the debt after the chasing calls than those with no contact information. The coefficient on $Chasing_{All} \times Contact_{Existence}$ is 1.17, significant at one percent level (t-statistic = 6.68), which accounts for an increase in the daily repayment likelihood of 41.1% relative to the sample mean in the window [1, 93].

Then, we decompose $Chasing_{All}$ into $Chasing_{Day\ 4,16}$ and $Chasing_{Day\ 31,61,91}$ to differentiate the first two rounds of chasing calls from the last three chasing calls. In Models 3 and 4, we find positive and significant coefficients on both interaction terms. For instance, the coefficients on $Chasing_{Day\ 4,16} \times Contact_{Existence}$ and $Chasing_{Day\ 31,61,91} \times Contact_{Existence}$ are 1.40 and 0.63 in Model 3, significant at one percent level (t-statistics = 6.46 and 4.89). In Models 3 and 4, the differences in coefficients are marginally significantly at ten percent level between $Chasing_{Day\ 4,16}$ and $Chasing_{Day\ 31,61,91}$ (z-statistics of differences = 3.11 and 2.04).

These findings imply that although chasing calls using digital information in the last three rounds are still effective in enforcing debt collection, the magnitude of the effect declines, consistent with our results in Table 4 for a comparison between the calls on Day 4 and Day 16. A potential explanation is that the delinquent borrowers, who have been warned using digital collateral but fail to make repayments after the first two chasing calls, are less concerned about their social capital potentially being damaged or being physically located. Therefore, these remaining delinquent borrowers are less responsive to further calls.

[Insert Table 8 Here]

3.5.3. Analyses based on alternative sample selections

For some delinquent borrowers in our treatment groups, the lender only has the digital footprint information about borrowers' frequent contacts or physical addresses. For other delinquent borrowers, the lender has acquired both the contact and address information from the third-party data providers.

In Table 9, we examine whether our main findings are driven by the borrowers having two types of digital footprints, or whether only one type of digital information matters for debt collection. In Model 1, we start with the 14,768 treatment borrowers with key contract information in the matched sample. We then exclude those with address information, which provides us with a sub-sample of 2,914 borrowers. We re-run the DID analyses with these *contract-only* borrowers and find that the coefficient on $Chasing_{Day\ 4} \times Contact_{Number}$ remains

positive and significant (coefficient = 2.57 and *t*-statistic = 3.63). The matched sample includes 34,968 borrower-day observations (2,914 × 2 × 6).

In Model 2, we find similar results when we perform the DID analyses in a sub-sample of 3,024 *address-only* borrowers (coefficient on $Chasing_{Day\ 4} \times Address_{Number}$ = 1.40 and *t*-statistic = 2.19). In this model, we follow a similar sample selection approach that leads a matched sample with 36,288 borrower-day observations (3,024 × 2 × 6). Finally, in Models 3 and 4, we investigate a sub-sample of 7,087 delinquent borrowers with digital footprints containing both key contact and physical address information. In a matched sample of 85,044 borrower-day observations (7,087 × 2 × 6), the coefficients on $Chasing_{Day\ 4} \times Contact_{Number}$ and $Chasing_{Day\ 4} \times Address_{Number}$ are 2.47 and 1.87 (*t*-statistics = 5.35 and 4.31), respectively.

These findings collectively suggest both types of digital footprints play significant roles in facilitating the debt collection process.

[Insert Table 9 Here]

4. Additional tests

In this section, we conduct the analyses for other outcome variables along the loan process, associated with the repayment amount, the likelihood of loan delinquency, and the decision of loan approval.

4.1. Repayment amount for delinquent loans

To this end, we mainly focus on the probability of repayment for delinquent loans. This sub-section will investigate the impact of digital collateral through chasing calls on the amount of debt recovery. This analysis can provide important implications for the lender. When a delinquent borrower makes loan repayment, she may underpay the overdue amount because of a shortfall of cash or make payment more than the original amount for the overdue penalty. Therefore, from the lender's perspective, the expected payoff for an average delinquent loan is a function of the likelihood of repayment, and the perceived amount of repayment contingent on the repayment is made.

We construct two measures for debt recovery, $Repayment_{Ratio}$, and $Repayment_{Complete}$. The former is the ratio of the amount of repayment for the delinquent loan divided by the amount required to be paid by the due date, and the latter is the indicator variable equal to one if a borrower makes the repayment equal to or greater than the delinquent amount. The analyses are performed in the following model in a sample of 41,711 delinquent borrowers:

$$Repayment_{Ratio} / Repayment_{Complete} = \alpha + \beta_{DC} Digital\ Collateral\ Proxy + \beta_{Control} Control\ Variables + \beta_{FE} Fixed\ Effects + \varepsilon, \quad (2)$$

where *Digital Collateral Proxy* and *Control Variables* denote the metrics of digital collateral and control variables, which have been adopted in Equation (1). *Fixed Effects* represent the province and year fixed effects to control for the cross-sectional variations across geographic areas and also the time-series variations. The standard errors are clustered at the province level. We expect to find results consistent with the analyses of the repayment likelihood, i.e., a significant and positive coefficient on *Digital Collateral Proxy*, suggesting that the amount of debt recovery also increases with the availability of digital collateral.

Table 10 presents the results for the tests on the amount of debt recovery. In Models 1 and 2, we find that the coefficients on $Contact_{Existence}$ and $Address_{Existence}$ are 2.37 and 6.12 for $Repayment_{Ratio}$ and significant at the one percent level (*t*-statistics = 5.50 and 7.19). In terms of economic significance, borrowers with frequent mobile contacts (physical addresses) have a higher ratio of repayment to total debt outstanding than those with no digital collateral, accounting for an increase of 2.21% (5.71%) relative to the sample mean. Regarding our control variables, we find the intuitive results confirming that borrowers with higher risk profiles make the lower amounts of repayments, and borrowers with larger balances in online payment accounts pay off more outstanding debts. For example, in Model 1, the coefficients on *Score* and *Wealth* are -9.81 and 1.01 and significant at the one percent level (*t*-statistics = -12.00 and 10.25), respectively.

Furthermore, we focus on the completion of repayment, $Repayment_{Complete}$, in Models 3 and 4. The results show that borrowers with frequent contact information and physical address information are more likely to pay off the delinquent amounts fully. The coefficients on $Contact_{Existence}$ is 1.22, and that on $Address_{Existence}$ is 5.21, both significant at one percent level (*t*-statistics = 3.49 and 9.67). Economically, relative to the sample means, borrowers with contact (address) information have a 1.65% (7.06%) higher probability of debt recovery than those with no such information.

These results suggest that the magnitude of debt recovery increases with the availability of borrowers' social network and physical location information in digital footprints.

[Insert Table 10 Here]

4.2. Loan delinquency likelihood

Another possible concern regarding our main findings is that digital collateral measures could also reflect borrowers' creditworthiness (Berg et al., 2020). We have largely alleviated such concern by performing the DID analyses and controlling for borrowers' characteristics, including the metrics of credit quality such as *Score*, *Wealth*, and *History*.

Nevertheless, in this sub-section, we intend to further mitigate this concern by investigating the association between digital collateral and the likelihood of loan delinquency. The tests will be performed in the model below with a sample, including 97,783 loan contracts for both delinquent and non-delinquent borrowers:

$$Delinquency = \alpha + \beta_{DC}\, Digital\ Collateral\ Proxy + \beta_{Control}\, Control\ Variables + \beta_{FE}\, Fixed\ Effects + \varepsilon, \quad (3)$$

where *Delinquency* is an indicator variable equal to one if a borrower fails to make a loan payment by the due date at least once during the loan term. The control variables and fixed effects are the same as specified in Equation (2).

If the digital collateral metrics mainly capture the effect of borrowers' credit quality, we would expect a significantly negative association between *Delinquency* and *Digital Collateral Proxy*. However, the results reported in Table 11 indicate that the incidence of loan delinquency is positively associated with the presence of borrowers' frequent contact information. That is, in Model 1, we find that the coefficient on $Contact_{Existence}$ is 3.85 and significant at the one percent level (*t*-statistic = 15.73). Economically, this coefficient accounts for 8.98% of the average likelihood of loan delinquency in our sample. We do not find a significant coefficient on $Address_{Existence}$ in Model 2 (coefficient = -0.848, and *t*-statistic = -1.65). These findings suggest that our digital footprint proxies cannot be considered as the borrowers' credit quality measures with respect to their delinquency probability.

Regarding the control variables, we find older and female borrowers less likely involve in delinquency, because older people may have built up a long-term credit record, and female borrowers could be less aggressive in making lending decisions delinquency (e.g., coefficient on *Age* = -8.99 and *t*-statistic = -7.15, and coefficient on *Gender* = -4.10 and *t*-statistic = -6.89 in Model 1). For other conventional creditworthiness measures, consistent with prior studies, we find that they are negatively associated with the loan delinquency rate (e.g., Berg et al., 2020). For instance, the coefficients on *Score* are significantly positive across all the models, aligned with the notion that risky borrowers tend to engage in loan delinquency (e.g., coefficient = 11.41, and *t*-statistic = 16.79 in Model 1). Borrowers with less wealth and more rejections in previous loan applications are more likely to experience delinquency. For example,

in Model 1, the coefficient on *Wealth* is -1.86 (*t*-statistic = -27.36), and that on *History* is 3.20 (*t*-statistic = 5.15).

Overall, the findings in this sub-section mitigate the concern that the digital footprint information mainly reflects the credit quality of borrowers rather than capturing the effect of digital collateral.

[Insert Table 11 Here]

4.3. Loan approval decision

We finally examine the effect of digital collateral on the decision of loan approval made by the fintech lender. The analyses are conducted in a sample of 236,967 loan applications in the following model:

$$Approval = \alpha + \beta_{DC}\, Digital\ Collateral\ Proxy + \beta_{Control}\, Control\ Variables + \beta_{FE}\, Fixed\ Effects + \varepsilon, \quad (4)$$

where *Approval* is an indicator variable equal to one if a borrower's loan application is approved by the fintech lender, and zero otherwise. We include the same control variables and the province and year fixed effects as in Equation (2).

In Table 12, we present the evidence indicating that the fintech lender is more likely to approve the loan application from a borrower with digital collateral acquired through digital footprint. For example, we find the coefficients on $Contact_{Existence}$ and $Address_{Existence}$ are 0.97 and 4.67, significant at the one percent level (*t*-statistics = 3.17 and 14.42). Concerning the economic magnitude, these results imply that borrowers with frequent mobile contacts (physical address information) have an increased approval likelihood of 2.21% (11.27%) higher than those with no digital collateral, compared to the sample mean.

Moreover, we also find that the coefficients on control variables are consistent with prior studies (e.g., Berg et al. 2020; Agarwal et al., 2020). A borrower with higher risk profiles (*Score*), less wealth in online balance (*Wealth*), and more rejections from other marketplace lending platforms (*History*), is less likely to be approved by the fintech lender for the loan application. For example, in Model 1, the coefficients on *Score*, *Wealth*, and *History* are -35.25, 3.22, and -9.26, and all significant at one percent level.

In sum, the results show that the lender does consider borrowers' digital collateral during the loan application procedure. This is possibly due to the fact documented in our main findings that digital collateral acquired from borrowers' digital footprints can be used in debt collection when borrowers fail to make repayment on time in the future due date.

[Insert Table 12 Here]

5. Further discussion

5.1. Generality and implication

Understanding the allocation of financial resources in the capital market is a fundamental question. Financial institutions traditionally rely on collateral and credit information (i.e., income and assets) to overcome default risk in the retail debt market. Given information asymmetry between the lender and individual borrowers (especially smaller borrowers), the collateral allows for hedging against potential moral hazardous behavior of borrowers. I In the event of a borrower becomes delinquent, the financial institution will be protected by collateral. The assets used as collateral will be seized and sold to recover debt claims. Thus, individuals without adequate collateral are difficult to secure a loan from conventional financial institutions. Under this situation, the asymmetric information and lack of collateral create moral hazard and adverse selection problems for financial institutions, resulting in the inefficiency of resource allocation (e.g., Gine et al., 2012; Lin et al., 2013; Hildebrand et al., 2017).

The uncollateralized nature of marketplace lending in China provides an alternative mechanism in allocating financial resources to those individuals without sound credit records and collateral. The marketplace lending is characterized by small-amount and unsecured loans, mainly targeting young subprime borrowers with short-term credit needs.[15] The critical question to ask here is: Given that loans are unsecured (i.e., assets as collateral are not needed and even information of credit score is not required), why would marketplace institutions lend the money to the individuals? How could the lenders protect their claims in case of loan delinquency?

Given the advance of information technology in recent years (i.e., cloud computing, big data, and social network), information searching and sharing through digital footprints start to be utilized by marketplace lenders. Specifically, technological advancement allows lenders to gather and interpret various borrowers' information. The information, for example, including: 1) personal data about credit risk scores, career history, and job salary; 2) social network and connections with key contacts, close friends, and family members; and 3) online behavior with regard to social media activities and online shopping records.

[15] It has been argued that conventional financial institutions often deny young people's loan applications due to short credit history (e.g., Agarwal et al., 2020).

In this paper, we exploit the setting of uncollateralized nature for individual loans issued by a fintech lender in China and investigate the effectiveness of the debt collection tactics where borrowers' personal information is targeted in digital footprints. This is motivated by the limited evidence on whether digital information of an individual can substitute for borrowers' properties and chattels to protect debt claims. We show that the utilization of digital footprints as collateral represents a key difference between marketplace lenders and traditional financial institutions, which helps lenders address moral hazards and improve the efficiency of financial resource allocation.

Our paper attempts to provide the first direct evidence on how digital footprints can serve as collateral, facilitating the enforcement of debt contracts for marketplace lenders.

5.2. Differentiation from concurrent research

Liao et al. (2020) is one recent paper close to our study. In a sample of approximately 2,300 loans, they find that when a lender makes phone calls to the contacts of delinquent borrowers, the targeted borrowers are more likely to experience an ultimate loan default, suggesting that the collection tactic can lead to a negative reciprocity effect. In contrast, using a large sample of nearly 42,000 delinquent contracts, we find a positive impact of digital collateral on debt recovery.

Another key difference between the two studies is the choice of debt collection strategy. The lender in Liao et al. (2020) has limited resources to make phone calls, and thus strategically chooses some of the borrowers to apply the debt collection tactic. The collection strategy is also applied in selective dates at the lender's discretion. On the contrary, the lender in our sample applies one collection strategy consistently to all the borrowers by making the phone calls on several given days after delinquency. Therefore, the debt collection strategy in our study is indiscriminately executed to every borrower at fixed dates, exogenous to the characteristics of loans and borrowers.

Moreover, different from Liao et al. (2020), the data in our study allows us to observe a borrower's repayment decision at each day after delinquency, so that we can better assess the borrower's behavior on a daily basis, closely tied to the timing of chasing calls. Finally, the digital footprint data in our study includes the information of physical location, which has not been studied in Liao et al. (2020) but can be useful to facilitate debt collection in chasing calls.

5.3. Differentiation from microfinance literature

In the literature on microfinance regarding the setting of joint liability group loans (Coleman, 1988; Karlan et al., 2009; Cassar et al., 2007), social collateral considers social capital as an essential element under joint liability mechanism. Such capital is about trust, group pressure, monitoring, and networking to encourage and enforce loan repayment (Cassar et al., 2007; Karlan and Valdivia, 2011; Lin et al., 2013).

By organizing borrowers into self-selected members of solidarity groups, social collateral effectively transfers the costly monitoring effort incurred by financial institutions onto the peer members (Geleta, 2014). Therefore, social capitals of uncollateralized borrowers are used as collateral to enhance their credibility on a collective basis. Should an individual in the group defaults, other members need to repay her outstanding debt or the whole group will lose access to future finance (e.g., Armendáriz de Aghion, 1999; Ghatak, 1999). Consequently, peer monitoring and pressure from members in the group (i.e., social collateral) can create a dual trust setting enforcing repayments (Barboza, 2019).

Also, under the joint liability mechanism, peers have access to private information (due to the previous relationship) that is not available to the lender and ex ante they are able to "distinguish between default due to moral hazard and default due to true negative personal shocks" (Karlan, 2007). Therefore, given that group lending creates an incentive for peers to screen, monitor and enforce other members in the group (Hermes and Lensink, 2007),[16] a borrower's decision of making repayment will build on a trust relationship representing mutual insurance under joint liability (Barboza, 2019). In summary, monitoring and pressure exerted by group leader encourage members' repayments and served as an insurance to cover the repayment in a collective credit agreement, and peer pressure acts as a guarantee for the loan recovery (Banerjeet et al., 1994; Al-Azzam et al., 2002).

It is important to note that, in our study, the use of digital footprints as collateral is different from the conceptual construct of social collateral in the microfinance literature. This is because, in our setting, the lender imposes a direct impact on a delinquent borrower via the potential damage of her social capital and the accessibility to her physical location, rather than relying on the mechanisms through peer monitoring and enforcement under joint liability documented in prior microfinance studies.

6. Conclusion

[16] These include social humiliation, public distrust within the community, altruism motives toward those members in the group, retaliation, and even the physical punishment, etc.

In this paper, we find that the information contained in borrowers' digital footprints can be used by lenders as collateral for future debt collection when borrowers fail to make repayments. The role of digital collateral in debt collection is operated through two potential channels, that is, the possible damages to borrowers' social capitals, and the accessibility to their physical locations. Our findings add to the fintech literature by documenting evidence that technological innovations can create value for innovators and adopters in the financial markets. We also complement the debt literature by showing that digital footprints, used as one type of non-physical collateral, can facilitate the collection of delinquent loans.

Appendix Variable Definitions

Variable	Definition
$Paid_{t}$	An indicator variable equal to one if a borrower makes the repayment on the Day *t*, and zero otherwise.
$Paid_{[i, j]}$	Cumulative repayment likelihood from Day *i* to Day *j* in percentage.
$Chasing_{Day\ t}$	An indicator variable equal to one for days in a three-day window [*t*, *t+2*], and zero otherwise.
$Chasing_{All}$	An indicator variable equal to one for days in windows [4, 6], [16, 18], [31, 33], [61, 63], and [91, 93], and zero otherwise.
$Chasing_{Day\ 4,16}$	An indicator variable equal to one for days in windows [4, 6], and [16, 18], and zero otherwise.
$Chasing_{Day\ 31,61,91}$	An indicator variable equal to one for days in windows [31, 33], [61, 63], and [91, 93], and zero otherwise.
$Repayment_{Ratio}$	Ratio of the amount of repayment for the delinquent loan divided by the amount required to be paid by the due date.
$Repayment_{Complete}$	An indicator variable equal to one if a borrower makes the repayment equal to or greater than the delinquent amount, and zero otherwise.
Delinquency	An indicator variable equal to one if a borrower fails to make loan payment by due date at least once during loan term, and zero otherwise.
Approval	An indicator variable equal to one if a borrower's local application is approved by the fintech lender, and zero otherwise.
$Contact_{Existence}$	An indicator variable equal to one if a borrower has at least one contact who has more than *ten* calls with borrower, and zero otherwise.
$Address_{Existence}$	An indicator variable equal to one if a borrower has at least one online shopping delivery address, and zero otherwise.
$Contact_{Number}$	Number of contacts who have more than *ten* calls with a borrower. Logarithm value is taken in regression analysis.
$Address_{Number}$	Number of a borrower's online shopping delivery addresses. Logarithm value is taken in regression analysis.
$Contact_{Pseudo}$	An indicator variable equal to one for a pseudo borrower assigned to a treatment group based on $Contact_{Existence}$, and zero otherwise.
$Address_{Pseudo}$	An indicator variable equal to one for a pseudo borrower assigned to a treatment group based on $Address_{Existence}$, and zero otherwise.
$Contact_{1\leq Number\leq 7}$	An indicator variable equal to one if a borrower has at least one and at most seven contacts who have more than *ten* calls with borrower, and zero otherwise.
$Contact_{Number>7}$	An indicator variable equal to one if a borrower has more than seven contacts with more than *ten* calls with borrower, and zero otherwise.
$Address_{Number=1}$	An indicator variable equal to one if a borrower has one online shopping delivery address, and zero otherwise.
$Address_{Number>1}$	An indicator variable equal to one if a borrower has more than one online shopping delivery addresses, and zero otherwise.

Appendix (Contunied)

Variable	Definition
Amount	Amount of loan principal in thousands of Chinese Yuan. Logarithm value is taken in regression analysis.
Rate	Interest rate of loan on annual basis.
Age	Age of a borrower in years. Logarithm value is taken in regression analysis.
Gender	An indicator variable equal to one if a borrower is female, and zero otherwise.
Score	Credit risk score provided by *Tencent*, with high value suggesting high risk profile. Logarithm value is taken in regression analysis.
Wealth	Balance of a borrower's *Alipay* account from *Alibaba* in Chinese Yuan. Logarithm value is taken in regression analysis.
History	Number of loan applications of a borrower rejected by other online lending platforms. Logarithm value is taken in regression analysis.

References

Agarwal, S., Alok, S., Ghosh, P., Gupta S., 2020. Financial Inclusion and Alternate Credit Scoring for the Millennials: Role of Big Data and Machine Learning in Fintech. Working Paper, National University of Singapore.

Aghion, P., Bolton, P., 1992. An Incomplete Contracts Approach to Financial Contracting. Review of Economic Studies 59, 473–494.

Al-Azzam, M., Hill, S., Sarangi, S., 2012. Repayment performance in group lending: Evidence from Jordan. Journal of Development Economics 97, 404-414.

Armendáriz de Aghion, B., 1999. On the design of a credit agreement with peer monitoring. Journal of Development Economics 60, 79-104.

Banerjee, A., Besley, T., Guinnane, T., 1994. Thy Neighbor's Keeper: The Design of a Credit Cooperative with Theory and a Test. The Quarterly Journal of Economics 109, 491-515.

Barboza, G., 2019. Social collateral, dynamic incentives, and joint liability in a two‐step trust game: Lending to the unbankable. Economic Notes 48, 12132.

Benmelech, E., Bergman, N.K., 2009. Collateral Pricing. Journal of Financial Economics 91, 339–360.

Berg, T., Burg, V., Gombovic, A., Puri, M., 2020. On the Rise of FinTechs: Credit Scoring Using Digital Footprints. Review of Financial Studies, Forthcoming.

Buchak, G., Matvos, G., Piskorski, T., Seru, A., 2018. Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks. Journal of Financial Economics 130, 453–483.

Bursztyn, L., Fiorin, S., Gottlieb, D., Kanz, M., 2019. Moral Incentives in Credit Card Debt Repayment: Evidence from a Field Experiment. Journal of Political Economy 127, 1641–1683.

Cadena, X., Schoar, A., 2012. Remembering to Pay? Reminders vs. Financial Incentives for Loan Payments. Working Paper, National Bureau of Economic Research.

Cassar, A., Crowley, L., Wydick, B., 2007. The effect of social capital on group loan repayment: Evidence from field experiments. The Economic Journal 117, 85-106.

Coleman, J.S., 1988. Social Capital in the Creation of Human Capital. American Journal of Sociology 94, 95–120.

DeMarzo, P.M., 2019. Presidential Address: Collateral and Commitment. Journal of Finance 74, 1587-1619.

Du, N., Li, L., Lu, T., Lu, X., 2020. Prosocial Compliance in P2P Lending: A Natural Field Experiment. Management Science 66, 315–333.

Fedaseyeu, V., 2020. Debt Collection Agencies and the Supply of Consumer Credit. Journal of Financial Economics, Forthcoming.

Fedaseyeu, V., Hunt, R.M., 2018. The Economics of Debt Collection: Enforcement of Consumer Credit Contracts. Working Paper, China Europe International Business School.

Ghatak, M., 1999. Group lending, local information and peer selection. Journal of Development Economics 60, 27-50.

Geleta, E.B., 2014. Social Capital as Collateral: Banking on the Poor. American Journal of Economics and Sociology 73, 108-125.

Gine, X., Goldberg, J., Yang, D., 2012. Credit market consequences of improved personal identification: field experimental evidence from Malawi. American Economic Review 102, 2923-2854.

Grennan, J., Michaely, R., 2020. FinTechs and the Market for Financial Analysis. Working Paper, Duke University.

Hart, O., Moore, J., 1994. A Theory of Debt Based on the Inalienability of Human Capital. Quarterly Journal of Economics 109, 841–879.

Hart, O., Moore, J., 1998. Default and Renegotiation: A Dynamic Model of Debt. Quarterly Journal of Economics 113, 1–41.

Hermes, N, Lensink, R., 2007. The empirics of microfinance: what do we know? The Economic Journal 117, 1-10.

Hildebrand, T., Puri, M., Rocholl, J., 2017. Adverse Incentives in Crowdfunding. Management Science 63, 587-900.

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

Jimenez, G., Salas, V., Saurina, J., 2006. Determinants of Collateral. Journal of Financial Economics 81, 255–281.

Jiang, J., Liao, L., Lu, X., Wang, Z., Xiang, H., 2019. Can Big Data Defeat Traditional Credit Rating? Working Paper, Tsinghua University.

Johnson, H., Stulz, R.M., 1985. An Analysis of Secured Debt. Journal of Financial Economics 14, 501–521.

Karlan, D., 2007. Social Connections and Group Banking. The Economic Journal 117, 52–84.

Karlan, D., Mobius, M., Rosenblat, T., Szeidl, A., 2009. Trust and Social Collateral. Quarterly Journal of Economics 124, 1307–1361.

Karlan, D., Valdivia, M., 2011. Teaching Entrepreneurship: Impact of business training on microfinance clients and institutions, Review of Economics and Statistics 93, 510-527.

Karlan, D., Morten, M., Zinman, J., 2015. A Personal Touch: Text Messaging for Loan Repayment. Working Paper, Yale University.

Kiyotaki, N., Moore, J., 1997. Credit Cycles. Journal of Political Economy 105, 211–248.

Liao, L., Wang, Z., Yan, H., Zhou, C., 2020. When FinTech Meets Privacy: The Consequence of Personal Information Misuse in Debt Collection. Working Paper, Tsinghua University.

Lin, M., Prabhala, N.R., Viswanathan, S., 2013. Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. Management Science 59, 17-35.

Lusardi, A., Mitchell, O.S., 2014. The Economic Importance of Financial Literacy: Theory and Evidence. Journal of Economic Literature 52, 5-44.

Medina, P.C., 2018. Selective Attention in Consumer Finance: Evidence from a Randomized Intervention in the Credit Card Market. Working Paper, Texas A&M University.

Putnam, R.D., 1995. Bowling Alone: America's Declining Social Capital. Journal of Democracy 6, 65–78.

Rampini, A.A., Viswanathan, S., 2010. Collateral, Risk Management, and the Distribution of Debt Capacity. Journal of Finance 65, 2293–2322.

Rampini, A.A., Viswanathan, S., 2013. Collateral and Capital Structure. Journal of Financial Economics 109, 466–492.

Schiantarelli, F., Stacchini, M., Strahan, P.E., 2020. Bank Quality, Judicial Efficiency, and Loan Repayment Delays in Italy. Journal of Finance, Forthcoming.

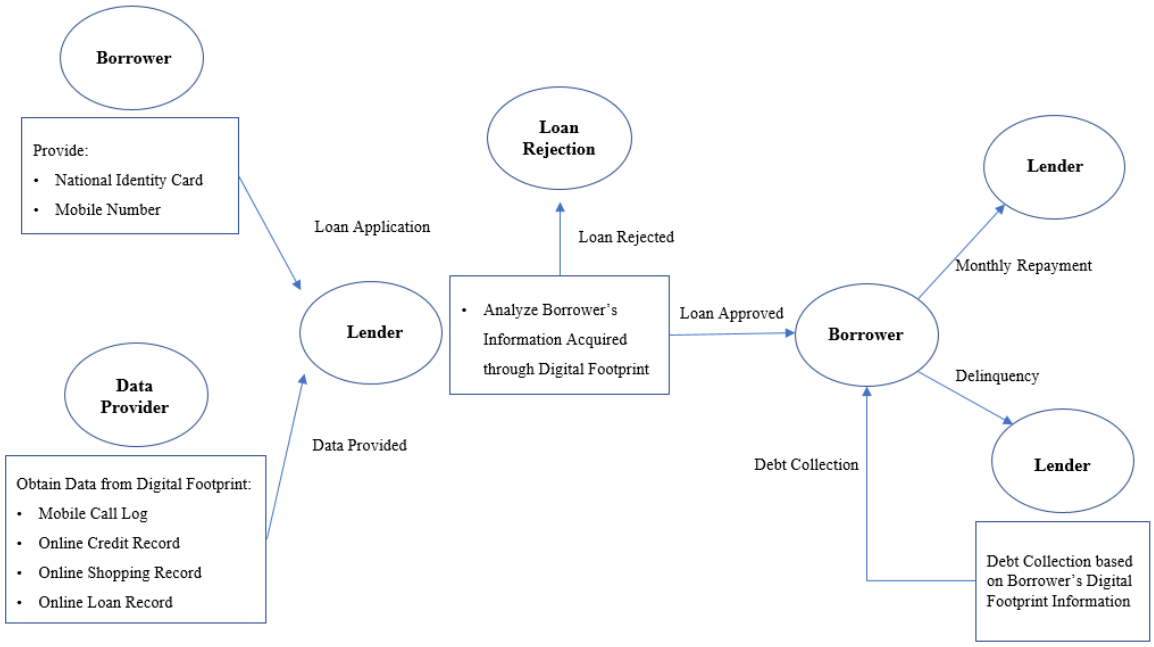
Tang, H., 2019. Peer-to-Peer Lenders Versus Banks: Substitutes or Complements? Review of Financial Studies 32, 1900–1938.

Thakor, A.V., 2020. Fintech and Banking: What Do We Know? Journal of Financial Intermediation, Forthcoming.

Vallée, B., Zeng, Y., 2019. Marketplace Lending: A New Banking Paradigm? Review of Financial Studies 32, 1939–1982.

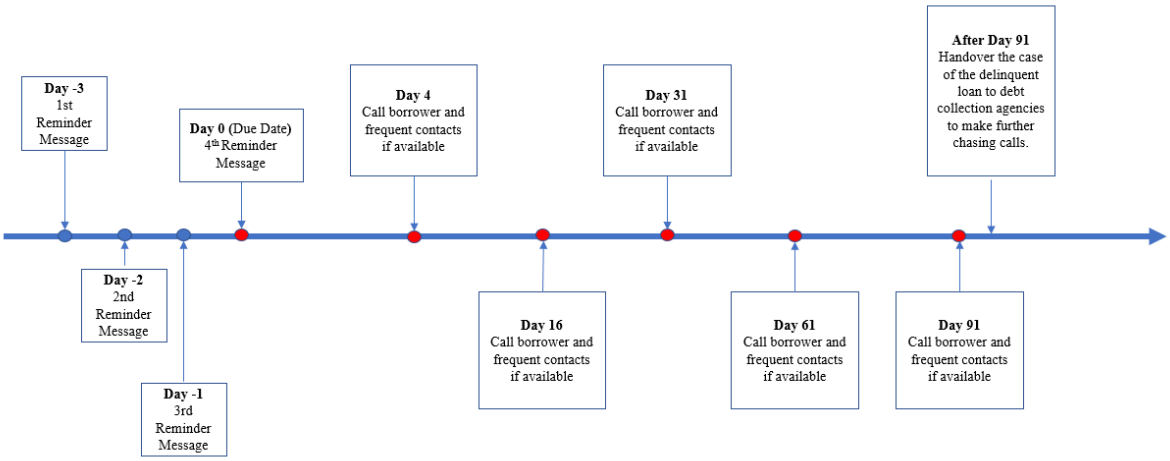
Zhu, C., 2019. Big Data as a Governance Mechanism. Review of Financial Studies 32, 2021–2061.

Figure 1 The Process of Loan Application, Approval, Repayment, and Collection



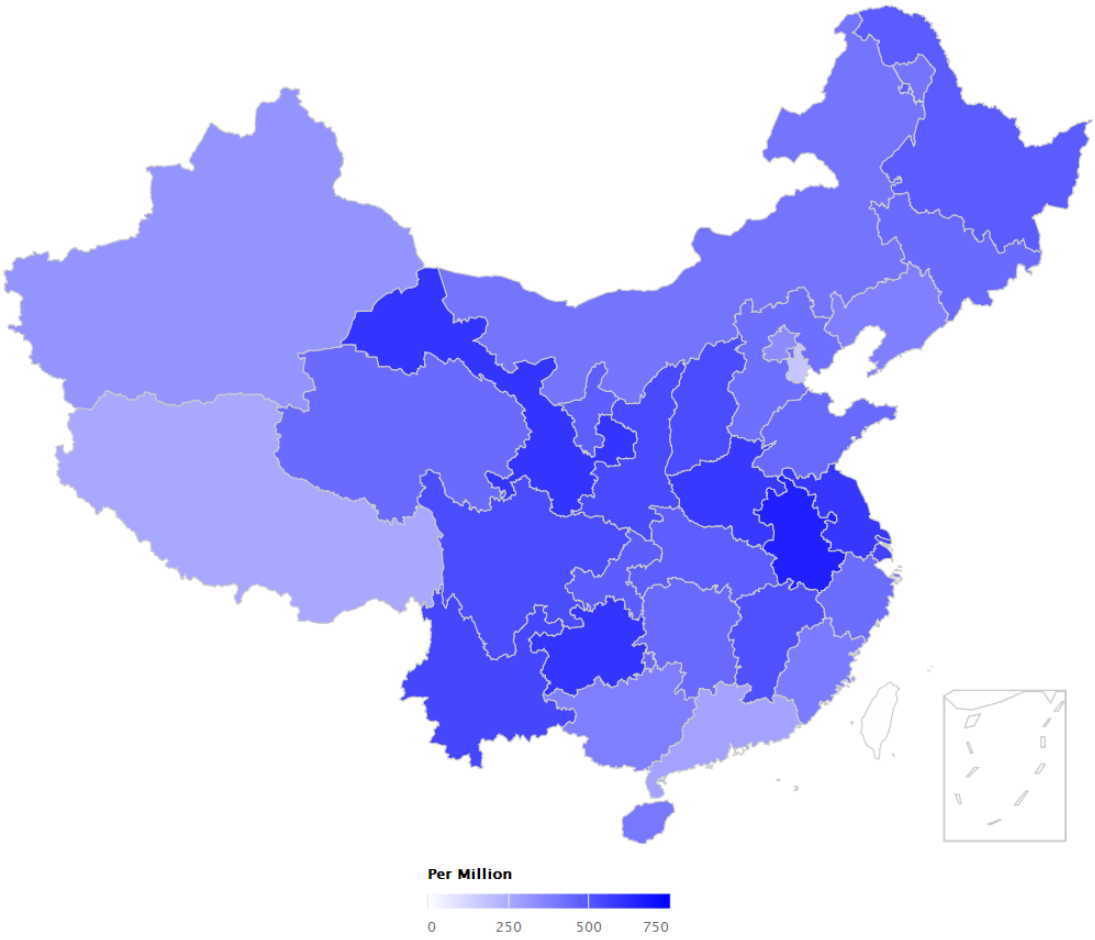
This figure illustrates the process of loan application, approval, repayment, and collection executed by the fintech lender. The sample analyzed in this paper includes 236,967 loan applications submitted to the fintech lender from July 2017 to November 2019, among which 97,783 loan applications have been approved and 41,711 approved loans experience borrower's delinquency.

Figure 2 Timeline of Debt Collection



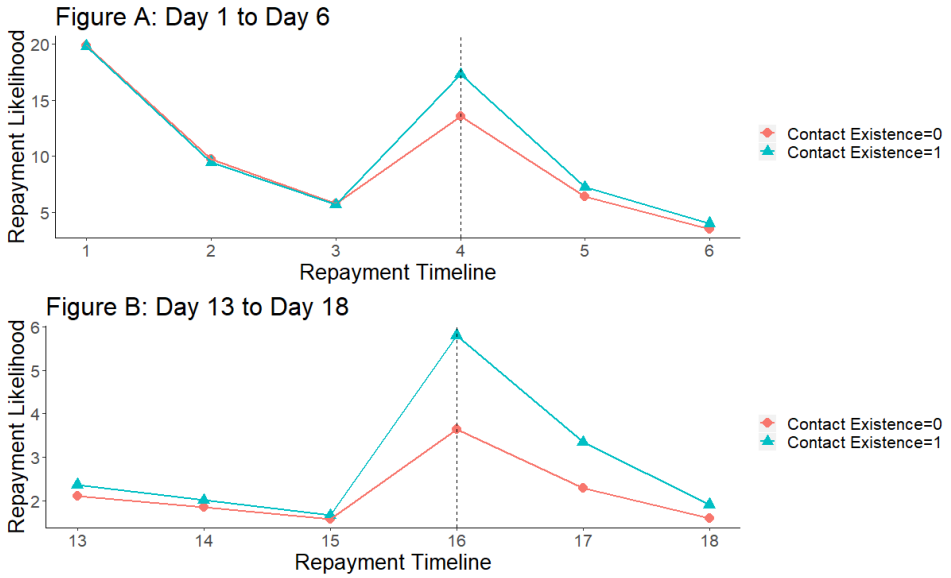
This figure presents the timeline of debt collection executed by the fintech lender, in which the Day *Zero* is the due date of month loan payment. The sample analyzed in this paper includes 97,783 approved loans from July 2017 to November 2019, among which 41,711 loans experience borrower's delinquency and enter the debt collection process.

Figure 3 Geographical Distribution of Loan Application



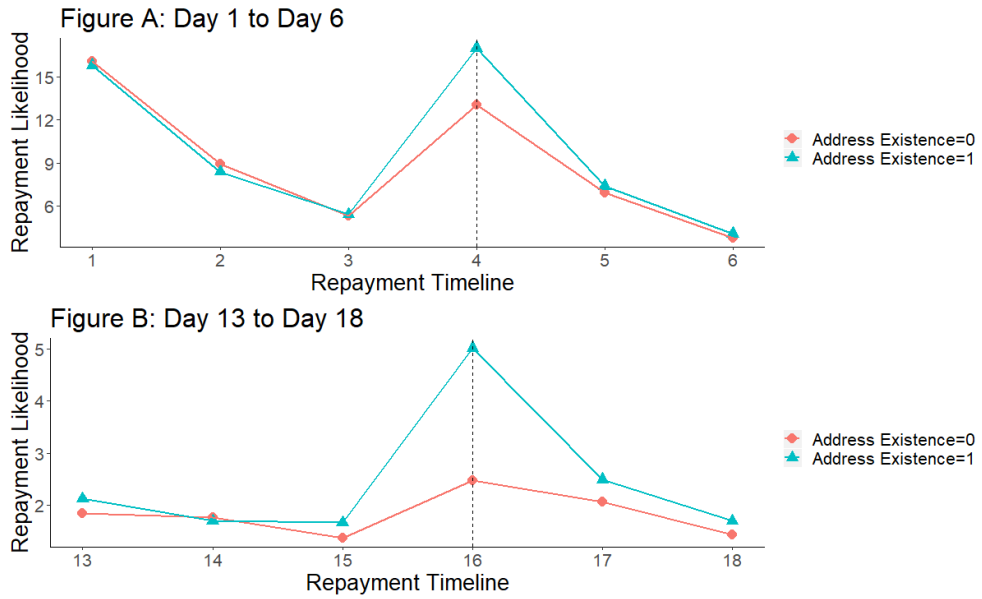
This figure plots the geographical distribution of loan applications submitted to the fintech lender. The color depth denotes the application density regarding the number of applications per million population in a province of China. The sample analyzed in this paper includes 236,967 loan applications submitted between July 2017 and November 2019, among which 97,783 loan applications have been approved.

Figure 4 Daily Repayment Likelihood around Chasing Calls based on Frequent Contact



This figure shows the daily likelihood of loan repayment around the first two rounds of chasing calls for the delinquent payment in Day 4 and Day 16. The due date is in Day *zero*. The repayment likelihoods of both treatment group and control group are presented after the propensity score matching, which is conducted based on the existence of borrower's frequent contact information. The full sample of delinquent borrowers analyzed in the analyses includes 41,711 delinquent loans.

Figure 5 Daily Repayment Likelihood around Chasing Calls based on Delivery Address



This figure shows the daily likelihood of loan repayment around the first two rounds of chasing calls for the delinquent payment in Day 4 and Day 16. The due date is in Day *zero*. The repayment likelihoods of both treatment group and control group are presented after the propensity score matching, which is conducted based on the existence of borrower's online shopping delivery address information. The full sample of delinquent borrowers analyzed in the analyses includes 41,711 delinquent loans.

Table 1 Summary Statistics

Variable	***OBS***	***Mean***	***STD***	***25%***	***Median***	***75%***
Paid $_{[1,6]}$	41,711	61.933	48.566	0	100	100
Paid $_{[1,18]}$	41,711	75.247	43.158	100	100	100
Paid $_{[1,93]}$	41,711	84.667	36.031	100	100	100
Repayment $_{Ratio}$	41,711	1.071	0.541	0.928	1.292	1.392
Repayment $_{Complete}$	41,711	0.738	0.440	0	1	1
Contact $_{Existence}$	41,711	0.651	0.476	0	1	1
Contact $_{Number}$	41,711	7.494	14.500	0	3	9
Address $_{Existence}$	41,711	0.765	0.423	0	1	1
Address $_{Number}$	41,711	1.06	0.808	0	1	1
Amount	41,711	3.966	1.297	3.000	3.980	4.780
Rate	41,711	0.295	0.083	0.205	0.357	0.359
Age	41,711	27.105	7.884	21	25	31
Gender	41,711	0.182	0.386	0	0	0
Score	41,711	56.761	15.261	48.000	60.200	66.750
Wealth	41,711	4.221	32.071	0.000	0.900	27.200
History	41,711	0.118	0.582	0	0	0
Delinquency	97,783	0.427	0.495	0	0	1
Approval	236,967	0.413	0.493	0	0	1

This table presents the summary statistics of the variables for the number of observations (*OBS*), mean (*Mean*), standard deviation (*STD*), the 25th (*25%*), median (*Median*), and 75th percentiles (*75%*) of the distributions of the variables. The samples of delinquent loans, approved loans, and loan applications consist of 41,711, 97,783, and 236,967 observations from July 2017 to November 2019, respectively. All variables are as defined in the Appendix.

Table 2 Treatment and Control Groups under Propensity Score Matching

Panel A: Propensity Score Matching based on *Contact* $_{Existence}$

	Contact $_{Existence}$ **= 1**	***Contact*** $_{Existence}$ **= 0**		
	Mean	***Mean***	***Difference***	***t-statistic***
Paid $_{[1,6]}$	63.591	58.830	4.761	8.421***
Paid $_{[1,18]}$	77.594	71.777	5.817	11.538***
Paid $_{[1,93]}$	86.301	81.827	4.474	10.313***
Amount	3.714	3.710	0.004	0.475
Rate	0.294	0.293	0.001	0.152
Age	27.611	27.679	-0.068	−0.858
Gender	0.175	0.174	0.001	0.068
Score	56.036	56.205	-0.169	−0.085
Wealth	6.080	6.203	-0.123	−0.982
History	0.046	0.045	0.001	0.947

Panel B: Propensity Score Matching based on *Address* $_{Existence}$

	Address $_{Existence}$ **= 1**	***Address*** $_{Existence}$ **= 0**		
	Mean	***Mean***	***Difference***	***t-statistic***
Paid $_{[1,6]}$	58.065	54.072	3.993	5.708***
Paid $_{[1,18]}$	72.297	67.681	4.616	7.149***
Paid $_{[1,93]}$	83.711	78.907	4.804	8.645***
Amount	3.565	3.601	-0.036	−1.562
Rate	0.313	0.308	0.005	1.787*
Age	31.083	30.487	0.596	1.222
Gender	0.171	0.170	0.001	0.131
Score	58.323	58.382	-0.059	−0.329
Wealth	1.804	2.056	-0.252	−6.386***
History	0.052	0.060	-0.008	−1.387

This table presents the comparison of the characteristics between the treatment group and control group after the propensity score matching. In Panel A, the treatment (control) group includes delinquent borrowers who have at least one frequent mobile contact (no frequent mobile contact) in digital footprint. In Panel B, the treatment (control) group includes delinquent borrowers who have at least one online shopping delivery address (no delivery address) in digital footprint. The samples of treatment and control groups both consist of 14,768 (9,994) delinquent loans from July 2017 to November 2019 in Panel A (Panel B). ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 3 Univariate Analyses of Repayment Likelihood

Panel A: Univariate Analyses based on *Contact* $_{Existence}$

	Contact $_{Existence}$ **= 1**	***Contact*** $_{Existence}$ **= 0**		
$Paid_t$	***Mean***	***Mean***	***Difference***	***t-statistic***
Day 1 to Day 3	11.683	11.780	−0.097	-0.449
Day 4 to Day 6	9.514	7.830	1.684	8.926***
Δ [4,6] - [1,3]	−2.169	−3.950	1.781	6.212***
Day 13 to Day 15	2.023	1.874	0.149	0.869
Day 16 to Day 18	3.681	2.520	1.161	5.387***
Δ [16,18] - [13,15]	1.658	0.646	1.012	3.722***

Panel B: Univariate Analyses based on *Address* $_{Existence}$

	Address $_{Existence}$ **= 1**	***Address*** $_{Existence}$ **= 0**		
$Paid_t$	***Mean***	***Mean***	***Difference***	***t-statistic***
Day 1 to Day 3	9.867	10.102	−0.235	-0.965
Day 4 to Day 6	9.488	7.921	1.567	6.828***
Δ [4,6] - [1,3]	−0.379	−2.181	1.802	5.378***
Day 13 to Day 15	1.834	1.663	0.171	0.936
Day 16 to Day 18	3.070	1.994	1.070	4.905***
Δ [16,18] - [13,15]	1.230	0.331	0.899	3.196***

The table presents the univariate analyses of the daily loan repayment likelihood between treatment group and control group after the propensity score matching. $Paid_t$ is an indicator variable equal to one if a borrower makes the repayment on the Day *t*, and zero otherwise. Day *Zero* is the due date of loan repayment. In Panel A, the treatment (control) group includes delinquent borrowers who have at least one frequent mobile contact (no frequent mobile contact) in digital footprint. In Panel B, the treatment (control) group includes delinquent borrowers who have at least one online shopping delivery address (no delivery address) in digital footprint. Around Day 4 (Day 16), the samples of treatment and control groups both consist of 14,768 and 9,994 (4,412 and 3,459) delinquent loans from July 2017 to November 2019 in Panel A and Panel B, respectively. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 4 Analyses of Repayment Likelihood around the First Two Rounds of Chasing Calls

Dependent Variable	***Paid*** $_t$			
	***T*=4**		***T*=16**	
	(1)	**(2)**	**(3)**	**(4)**
Chasing $_{Day\ T}$ × *Contact* $_{Existence}$	1.759***		1.033***	
	(5.805)		(3.666)	
Chasing $_{Day\ T}$ × *Address* $_{Existence}$		1.794***		0.898***
		(5.002)		(3.037)
Chasing $_{Day\ T}$ × *Amount*	−1.919***	−1.539***	0.529	0.600*
	(−4.618)	(−3.432)	(0.359)	(1.695)
Chasing $_{Day\ T}$ × *Rate*	−12.740***	−17.158***	−2.485	−3.527**
	(−6.976)	(−8.254)	(−1.477)	(−2.010)
Chasing $_{Day\ T}$ × *Age*	5.710***	5.142***	−1.137*	−0.252
	(10.211)	(8.407)	(−2.410)	(−0.524)
Chasing $_{Day\ T}$ × *Gender*	−2.289***	−1.243***	−0.031	−0.354
	(−5.722)	(−2.587)	(−0.079)	(−0.854)
Chasing $_{Day\ T}$ × *Score*	3.349***	3.210***	-0.329	-0.027
	(6.497)	(4.751)	(-0.558)	(-0.038)
Chasing $_{Day\ T}$ × *Wealth*	−0.327***	−0.158	0.089	−0.041
	(−5.119)	(−1.223)	(1.434)	(−0.342)
Chasing $_{Day\ T}$ × *History*	−0.553	−0.743	0.566	0.072
	(−0.745)	(−0.928)	(0.832)	(0.099)
Borrower Fixed Effect	Yes	Yes	Yes	Yes
Day Fixed Effect	Yes	Yes	Yes	Yes
Observations	177,216	119,928	52,944	41,508
R^2_{ADJ}	10.9%	10.8%	1.8%	1.8%

This table presents the difference-in-difference analyses of the daily loan repayment likelihood based on lender's first two rounds of chasing calls made on Day 4 and Day 16. The analyses are performed on the delinquent-borrower-day level. The dependent variable, *Paid* $_t$, is an indicator variable equal to one if a borrower makes the repayment on the Day *t*, and zero otherwise. Day *Zero* is the due date of loan repayment. In Models 1 and 2, the tests are conducted in window [1, 6], where Day 4 is the date that the lender makes the first round of chasing calls for delinquent loans. In Models 3 and 4, the tests are conducted in window [13, 18], where Day 16 is the date that the lender makes the second round of chasing calls for delinquent loans. *Chasing* $_{Day\ T}$ is an indicator variable equal to one for days in window [*T*, *T+2*], and zero for window [*T-3*, *T-1*]. The dependent variable, *Paid* $_t$, is regressed on *Chasing*, *Chasing* × *Digital Collateral Proxy* (i.e., *Contact* $_{Existence}$ or *Address* $_{Existence}$), controlling for the interactions between the *Chasing* dummies and other characteristics of loan and borrower, as well as the borrower and day fixed effects. In Models 1 and 2 (Models 3 and 4), the matched samples based on *Contact* $_{Existence}$ and *Address* $_{Existence}$ consist of 177,216 and 119,928 (52,944 and 41,508) borrower-day observations from July 2017 to November 2019, respectively. The *t*-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and borrower-level clustering. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 5 Placebo Tests Based on Pseudo Chasing Call Date

Dependent Variable	***Paid*** $_t$	
	(1)	**(2)**
Chasing $_{Day\ 1}$ × *Contact* $_{Existence}$	−0.074	
	(−0.400)	
Chasing $_{Day\ 1}$ × *Address* $_{Existence}$		−0.266
		(−1.220)
Chasing $_{Day\ 1}$ × *Amount*	1.074***	0.956***
	(4.219)	(3.507)
Chasing $_{Day\ 1}$ × *Rate*	5.977***	10.981***
	(5.281)	(8.508)
Chasing $_{Day\ 1}$ × *Age*	−4.664***	−3.804***
	(−13.662)	(−10.232)
Chasing $_{Day\ 1}$ × *Gender*	1.566***	1.048***
	(6.312)	(3.559)
Chasing $_{Day\ 1}$ × *Score*	−4.144***	−4.014***
	(−13.515)	(−9.839)
Chasing $_{Day\ 1}$ × *Wealth*	0.419***	0.305***
	(10.831)	(3.934)
Chasing $_{Day\ 1}$ × *History*	0.795*	1.367***
	(1.791)	(2.902)
Borrower Fixed Effect	Yes	Yes
Day Fixed Effect	Yes	Yes
Observations	177,216	119,928
R^2_{ADJ}	5.6%	4.2%

This table presents the placebo tests of the daily loan repayment likelihood in a six-day window [-2, 3] with no chasing call. The analyses are performed on the delinquent-borrower-day level. The dependent variable, *Paid* $_t$, is an indicator variable equal to one if a borrower makes the repayment on the Day *t*, and zero otherwise. Day *Zero* is the due date of loan repayment. *Chasing* $_{Day\ 1}$ is an indicator variable equal to one for days in window [1, 3], and zero for window [-2, 0]. *Paid* $_t$ is regressed on *Chasing*, *Chasing* × *Digital Collateral Proxy* (i.e., *Contact* $_{Existence}$ or *Address* $_{Existence}$), controlling for the interactions between the *Chasing* dummies and other characteristics of loan and borrower, as well as the borrower and day fixed effects. The matched samples based on *Contact* $_{Existence}$ (*Address* $_{Existence}$) consist of 177,216 (119,928) borrower-day observations from July 2017 to November 2019. The *t*-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and borrower-level clustering. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 6 Placebo Tests Based on Pseudo Borrowers with Digital Collateral

Dependent Variable	***Paid*** $_t$	
	(1)	**(2)**
Chasing $_{Day\ 4}$ × *Contact* $_{Pseudo}$	−0.467	
	(−1.128)	
Chasing $_{Day\ 4}$ × *Address* $_{Pseudo}$		-0.155
		(−0.322)
Chasing $_{Day\ 4}$ × *Amount*	-2.385***	-1.349**
	(-4.276)	(-2.285)
Chasing $_{Day\ 4}$ × *Rate*	-15.333***	-25.041***
	(-6.057)	(-9.429)
Chasing $_{Day\ 4}$ × *Age*	5.049***	6.331***
	(6.440)	(7.989)
Chasing $_{Day\ 4}$ × *Gender*	-2.385***	-1.278**
	(-4.276)	(-1.925)
Chasing $_{Day\ 4}$ × *Score*	4.360***	4.634***
	(6.057)	(5.038)
Chasing $_{Day\ 4}$ × *Wealth*	-0.405***	0.263
	(-4.681)	(1.502)
Chasing $_{Day\ 4}$ × *History*	-0.510	-2.093*
	(-0.512)	(-1.913)
Borrower Fixed Effect	Yes	Yes
Day Fixed Effect	Yes	Yes
Observations	88,608	59,964
R^2_{ADJ}	6.1%	6.4%

This table presents the placebo tests of the daily loan repayment likelihood in a six-day window [1, 6] for pseudo borrowers with digital collateral. The analyses are performed on the delinquent-borrower-day level. The dependent variable, *Paid* $_t$, is an indicator variable equal to one if a borrower makes the repayment on the Day *t*, and zero otherwise. Day *Zero* is the due date of loan repayment. *Chasing* $_{Day\ 4}$ is an indicator variable equal to one for days in window [4, 6], and zero for window [1, 3]. *Paid* $_t$ is regressed on *Chasing*, *Chasing* × *Pseudo Digital Collateral Proxy* (i.e., *Contact* $_{Pseudo}$ or *Address* $_{Pseudo}$), controlling for the interactions between the *Chasing* dummies and other characteristics of loan and borrower, as well as the borrower and day fixed effects. The matched samples based on *Contact* $_{Pseudo}$ (*Address* $_{Pseudo}$) consist of 88,608 (59,964) borrower-day observations from July 2017 to November 2019. The *t*-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and borrower-level clustering. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 7 Analyses of Contact and Address Numbers

Dependent Variable	***Paid*** $_t$			
	(1)	**(2)**	**(3)**	**(4)**
Chasing $_{Day\ 4}$ × *Contact* $_{Number}$	0.514***			
	(3.931)			
Chasing $_{Day\ 4}$ × *Address* $_{Number}$		1.843***		
		(4.425)		
Chasing $_{Day\ 4}$ × *Contact* $_{1\leq Number\leq 7}$			2.134***	
			(5.961)	
Chasing $_{Day\ 4}$ × *Contact* $_{Number>7}$			1.312***	
			(3.302)	
Chasing $_{Day\ 4}$ × *Address* $_{Number=1}$				1.846***
				(4.737)
Chasing $_{Day\ 4}$ × *Address* $_{Number>1}$				1.652***
				(2.891)
Chasing $_{Day\ 4}$ × *Amount*	−1.941***	−1.537***	−1.941***	−1.537***
	(−4.668)	(−3.429)	(−4.668)	(−3.429)
Chasing $_{Day\ 4}$ × *Rate*	−12.903***	−17.230***	−12.903***	−17.230***
	(−7.063)	(−8.287)	(−7.063)	(−8.287)
Chasing $_{Day\ 4}$ × *Age*	5.495***	5.236***	5.495***	5.236***
	(9.784)	(8.545)	(9.784)	(8.545)
Chasing $_{Day\ 4}$ × *Gender*	−2.253***	−1.251***	−2.253***	−1.251***
	(−5.628)	(−2.602)	(−5.628)	(−2.602)
Chasing $_{Day\ 4}$ × *Score*	3.366***	3.230***	3.366***	3.230***
	(6.525)	(4.781)	(6.525)	(4.781)
Chasing $_{Day\ 4}$ × *Wealth*	−0.326***	−0.157	−0.326***	−0.157
	(−5.093)	(−1.215)	(−5.093)	(−1.215)
Chasing $_{Day\ 4}$ × *History*	−0.631	−0.735	−0.631	−0.735
	(−0.850)	(−0.916)	(−0.850)	(−0.916)
Borrower Fixed Effect	Yes	Yes	Yes	Yes
Day Fixed Effect	Yes	Yes	Yes	Yes
Observations	177,216	119,928	177,216	119,928
R^2_{ADJ}	10.9%	10.8%	10.9%	10.8%

This table presents the analyses based on the amount of digital information. The analyses are performed on the delinquent-borrower-day level. The dependent variable, *Paid* $_t$, is an indicator variable equal to one if a borrower makes the repayment on the Day *t*, and zero otherwise. Day *Zero* is the due date of loan repayment. Day 4 is the date that the lender makes the first round of chasing calls for delinquent loans. *Chasing* $_{Day\ 4}$ is an indicator variable equal to one for days in window [4, 6], and zero for window [1, 3]. In Models 1 and 2, the dependent variables are the number of borrowers' frequent contacts and physical addresses, *Contact* $_{Number}$ and *Address* $_{Number}$, respectively. In Models 3 and 4, we construct four indicator variables equal to one for borrowers with contact number between one and seven (*Contact* $_{1\leq Number\leq 7}$), and greater than seven (*Contact* $_{Number>7}$), and for borrowers with address number equal to one (*Address* $_{Number=1}$), and greater than one (*Address* $_{Number>1}$), and otherwise zero. The matched samples based on *Contact* $_{Existence}$ (*Address* $_{Existence}$) consist of 177,216 (119,928) borrower-day observations from July 2017 to November 2019. The *t*-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and borrower-level clustering. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 8 Analyses of Further Chasing Calls

Dependent Variable	***Paid*** $_t$			
	All Chasing Calls		**Early versus Later Calls**	
	(1)	**(2)**	**(3)**	**(4)**
Chasing $_{All}$ × *Contact* $_{Existence}$	1.170***			
	(6.683)			
Chasing $_{All}$ × *Address* $_{Existence}$		1.188***		
		(6.151)		
Chasing $_{Day\ 4,16}$ × *Contact* $_{Existence}$			1.401***	
			(6.461)	
Chasing $_{Day\ 31,61,91}$ × *Contact* $_{Existence}$			0.631***	
			(4.893)	
Chasing $_{Day\ 4,16}$ × *Address* $_{Existence}$				1.376***
				(5.643)
Chasing $_{Day\ 31,61,91}$ × *Address* $_{Existence}$				0.804***
				(5.417)
Chasing $_{All}$ × *Amount*	−1.054***	−0.801**	−1.049***	−0.798***
	(−4.484)	(−3.371)	(−4.465)	(−3.357)
Chasing $_{All}$ × *Rate*	−6.745***	−8.850***	−6.725***	−8.827***
	(−6.431)	(−7.848)	(−6.415)	(−7.832)
Chasing $_{All}$ × *Age*	2.571***	2.475***	2.578***	2.447***
	(8.583)	(7.821)	(8.610)	(7.830)
Chasing $_{All}$ × *Gender*	−1.349***	-0.699***	−1.349***	−0.695***
	(−5.707)	(−2.644)	(−5.717)	(−2.630)
Chasing $_{All}$ × *Score*	2.148***	2.046***	2.141***	2.040***
	(6.207)	(4.725)	(6.186)	(4.713)
Chasing $_{All}$ × *Wealth*	-0.166***	-0.057	-0.165***	-0.057
	(-4.248)	(-0.786)	(-4.242)	(-0.773)
Chasing $_{All}$ × *History*	-0.157	-0.400	-0.163	-0.401
	(-0.361)	(0.866)	(-0.373)	(-0.869)
Borrower Fixed Effect	Yes	Yes	Yes	Yes
Day Fixed Effect	Yes	Yes	Yes	Yes
Observations	330,624	241,812	330,624	241,812
R^2_{ADJ}	5.0%	5.4%	5.0%	5.4%

This table presents the analyses of the daily loan repayment likelihood based on all the chasing calls made on Day 4, Day 16, Day 31, Day 61, and Day 91. The analyses are performed on the delinquent-borrower-day level. The dependent variable, *Paid* $_t$, is an indicator variable equal to one if a borrower makes the repayment on the Day *t*, and zero otherwise. Day *Zero* is the due date of loan repayment. The tests are conducted in five six-day windows [1, 6], [13, 18], [28, 33], [58, 63], and [88, 93] after the due date. In Models 1 and 2, *Chasing* $_{All}$ is an indicator variable equal to one for days in windows [4, 6], [16, 18], [31, 33], [61, 63], and [91, 93], and zero otherwise. In Models 3 and 4, *Chasing* $_{Day\ 4,16}$ is an indicator variable equal to one for days in windows [4, 6] and [16, 18], and zero otherwise, while *Chasing* $_{Day\ 31,61,91}$ is an indicator variable equal to one for days in windows [31, 33], [61, 63], and [91, 93], and zero otherwise. *Paid* $_t$ is regressed on *Chasing*, *Chasing* × *Digital Collateral Proxy* (i.e., *Contact* $_{Existence}$ or *Address* $_{Existence}$), controlling for the interactions between the *Chasing* dummies and other characteristics of loan and borrower, as well as the borrower and day fixed effects. The matched samples based on *Contact* $_{Existence}$ (*Address* $_{Existence}$) consist of 330,624 (241,812) borrower-day observations from July 2017 to November 2019. The *t*-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and borrower-level clustering. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 9 Analyses of Borrowers with Both or Either Contact and/or Address Information

Dependent Variable	***Paid*** $_t$			
	Non-Overlapping		**Overlapping**	
	(1)	**(2)**	**(3)**	**(4)**
Chasing $_{Day\ 4}$ × *Contact* $_{Number}$	2.567***		2.465***	
	(3.625)		(5.350)	
Chasing $_{Day\ 4}$ × *Address* $_{Number}$		1.397**		1.865***
		(2.186)		(4.314)
Chasing $_{Day\ 4}$ × *Amount*	-1.537*	−1.416***	-2.087***	-1.444***
	(−1.843)	(−1.708)	(-3.538)	(-2.674)
Chasing $_{Day\ 4}$ × *Rate*	−19.441***	−19.945***	-9.125***	-18.231***
	(−5.216)	(−5.322)	(-3.380)	(-7.253)
Chasing $_{Day\ 4}$ × *Age*	4.992***	3.739***	5.236***	5.334***
	(4.436)	(3.244)	(6.410)	(6.829)
Chasing $_{Day\ 4}$ × *Gender*	−1.918**	−1.614***	-1.712***	-1.484***
	(−2.416)	(−1.868)	(-2.977)	(-2.596)
Chasing $_{Day\ 4}$ × *Score*	2.953**	2.752**	3.150***	3.144***
	(2.490)	(2.353)	(4.145)	(3.973)
Chasing $_{Day\ 4}$ × *Wealth*	-0.213	0.043	-0.363***	-0.057
	(-1.283)	(0.191)	(-3.406)	(-0.362)
Chasing $_{Day\ 4}$ × *History*	-0.075	-0.383	0.283	-0.481
	(-0.037)	(-0.233)	(0.306)	(-0.514)
Borrower Fixed Effect	Yes	Yes	Yes	Yes
Day Fixed Effect	Yes	Yes	Yes	Yes
Observations	34,968	36,288	85,044	85,044
R^2_{ADJ}	6.1%	6.2%	6.8%	6.7%

This table presents the analyses of the daily loan repayment likelihood for subsample with no overlapping or fully overlapping on *Contact* $_{Existence}$ and *Address* $_{Existence}$, based on all the chasing call made on Day 4. The analyses are performed on the delinquent-borrower-day level. The dependent variable, *Paid* $_t$, is an indicator variable equal to one if a borrower makes the repayment on the Day *t*, and zero otherwise. Day *Zero* is the due date of loan repayment. The tests are conducted in the full window [1, 6] after the due date. In Models 1 (Model 2), the sample includes borrowers in treatment group with only the information of *Contact* $_{Number}$ (*Address* $_{Number}$). Models 3 and 4 include borrowers in treatment group having the information for both *Contact* $_{Number}$ and *Address* $_{Number}$. *Chasing* $_{Day\ 4}$ is an indicator variable equal to one for days in windows [4, 6], and zero otherwise. *Paid* $_t$ is regressed on *Chasing*, *Chasing* × *Digital Collateral Proxy* (i.e., *Contact* $_{Existence}$ or *Address* $_{Existence}$), controlling for the interactions between the *Chasing* dummies and other characteristics of loan and borrower, as well as the borrower and day fixed effects. In Models 1 and 2, the matched samples consist of 34,968 and 36,288 borrower-day observations, respectively, while in Models 3 and 4, the matched sample consists of 85,044 borrower-day observations from July 2017 to November 2019. The *t*-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and borrower-level clustering. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 10 Analyses of Repayment Amount for Delinquent Loan

Dependent Variable	***Repayment*** $_{Ratio}$		***Repayment*** $_{Complete}$	
	(1)	**(2)**	**(3)**	**(4)**
Contact $_{Existence}$	2.367***		1.223***	
	(5.498)		(3.487)	
Address $_{Existence}$		6.117***		5.208***
		(7.192)		(9.672)
Amount	−7.523***	−7.711***	−5.964***	−6.133***
	(−10.709)	(−10.829)	(−11.357)	(−11.276)
Rate	11.524***	11.085***	−11.760***	−12.261***
	(3.061)	(2.860)	(−3.430)	(−3.538)
Age	−23.668***	−21.800***	−16.669***	−15.135***
	(−11.834)	(−11.206)	(−11.379)	(−10.272)
Gender	4.683***	4.522***	5.188***	5.044***
	(6.581)	(6.478)	(9.722)	(9.775)
Score	−9.806***	−9.338***	−8.765***	−8.342***
	(−12.002)	(−12.063)	(−11.386)	(−11.361)
Wealth	1.005***	0.780***	1.038***	0.843***
	(10.248)	(8.880)	(14.487)	(11.315)
History	0.880	0.781	−0.009	−0.200
	(0.967)	(0.880)	(−0.012)	(−0.268)
Province Fixed Effect	Yes	Yes	Yes	Yes
Year Quarter Fixed Effect	Yes	Yes	Yes	Yes
Observations	41,711	41,711	41,711	41,711
R^2_{ADJ}	3.8%	3.9%	2.9%	3.1%

This table presents the analyses of the repayment amount for the delinquent loan. The analyses are performed on the loan level. In Models 1 and 2, the dependent variable, *Repayment* $_{Ratio}$, is the ratio for the amount of repayment for the delinquent loan divided by the amount required to be paid by the due date. In Models 3 and 4, the dependent variable, *Repayment* $_{Complete}$, is an indicator variable equal to one if a borrower makes the repayment for the amount equal to or greater than the delinquent amount, and zero otherwise. *Repayment* $_{Ratio}$ and *Repayment* $_{Complete}$ are regressed on the *Digital Collateral Proxy* (i.e., *Contact* $_{Existence}$, or *Address* $_{Number}$), controlling for other characteristics of loan and borrower, as well as the province and year fixed effects. The sample of delinquent loans consists of 41,711 observations from July 2017 to November 2019. The *t*-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and province-level clustering. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 11 Analyses of Loan Delinquency Likelihood

Dependent Variable	***Delinquency***	
	(1)	**(2)**
Contact $_{Existence}$	3.851***	
	(15.732)	
Address $_{Existence}$		−0.848
		(−1.652)
Amount	5.089***	5.124***
	(7.788)	(7.853)
Rate	13.355***	13.959***
	(6.377)	(6.862)
Age	−8.993***	−8.918***
	(−7.152)	(−6.688)
Gender	−4.097***	−4.042***
	(−6.892)	(−6.902)
Score	11.412***	11.223***
	(16.788)	(17.010)
Wealth	−1.857***	−1.807***
	(−27.360)	(−28.662)
History	3.195***	3.683***
	(5.152)	(6.218)
Province Fixed Effect	Yes	Yes
Year Quarter Fixed Effect	Yes	Yes
Observations	97,783	97,783
R^2_{ADJ}	6.1%	6.0%

This table presents the analyses of the loan delinquency likelihood. The analyses are performed on the loan level. The dependent variable, *Delinquency*, is an indicator variable equal to one if a borrower fails to make the monthly loan payment by the due date for at least one time during the loan term, and zero otherwise. *Delinquency* is regressed on the *Digital Collateral Proxy* (i.e., *Contact* $_{Existence}$, or *Address* $_{Number}$), controlling for other characteristics of loan and borrower, as well as the province and year fixed effects. The sample of loans consists of 97,783 observations from July 2017 to November 2019. The *t*-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and province-level clustering. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.

Table 12 Analyses of Loan Approval

Dependent Variable	***Approval***	
	(1)	**(2)**
Contact $_{Existence}$	0.974***	
	(3.167)	
Address $_{Existence}$		4.666***
		(14.423)
Amount	−14.994***	−15.119***
	(−20.758)	(−20.931)
Rate	107.984***	107.514***
	(18.402)	(18.538)
Age	−7.820***	−6.317***
	(−7.952)	(−6.291)
Gender	12.066***	11.950***
	(27.097)	(27.367)
Score	−35.251***	−34.888***
	(−95.932)	(−97.897)
Wealth	3.222***	3.014***
	(83.566)	(82.717)
History	−9.259***	−9.482***
	(−13.538)	(−14.469)
Province Fixed Effect	Yes	Yes
Year Quarter Fixed Effect	Yes	Yes
Observations	236,967	236,967
R^2_{ADJ}	13.2%	13.3%

This table presents the analyses of the loan approval likelihood. The analyses are performed on the loan application level. The dependent variable, *Approval*, is an indicator variable equal to one if a borrower's local application is approved by the fintech lender, and zero otherwise. *Approval* is regressed on the *Digital Collateral Proxy* (i.e., *Contact* $_{Existence}$, or *Address* $_{Number}$), controlling for other characteristics of loan and borrower, as well as the province and year fixed effects. The sample of loan applications consists of 236,967 observations from July 2017 to November 2019. The *t*-statistics shown in parentheses are based on standard errors adjusted for heteroskedasticity and province-level clustering. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All the variables are defined in the Appendix.