Cashless Payment and Financial Inclusion

Shumiao Ouyang*
Princeton University

March 2023

Click here for most recent version.

Abstract

This paper investigates how cashless payment affects credit access for the underprivileged using Alipay, a BigTech platform that offers various financial services to over 1 billion users. Leveraging a natural experiment and a representative Alipay user sample, I find that cashless payment adoption increases credit access by 56.3% and a 1% rise in payment flow increases credit line by 0.41%. These effects are stronger for the less educated and the older. Counterfactual analysis shows that cashless payment data increase credit lines by 57.7%, consumer surplus by 0.5% of median income, and lender profit by 41.3% of consumer surplus.

Keywords: Cashless Payment, BigTech, Consumer Credit, Financial Inclusion, Technology Adoption
JEL Codes: G21, G23, G51, G53, O33

I am grateful to Wei Xiong, Markus K. Brunnermeier, Atif R. Mian, and Ravi Jagannathan for their continuous guidance and support. For useful comments I also thank Mark Aguiar, Torben Andersen, Cynthia Mei Balloch, Paul Beaumont, James Brown, Greg Buchak, Long Chen, Lin William Cong, Natalie Cox, Kim Fe Cramer, Marco Di Maggio, John Driscoll, Elena Falcettoni, Hanming Fang, Andreas Fuster, Leonardo Gambacorta, Sean Higgins, Jing Huang, Sun Yong Kim, Faizaan Kisat, Nobuhiro Kiyotaki, Robert Korajczyk, Kwok Hao Lee, Moritz Lenel, Chen Lin, Ernest Liu, Tong Liu, Xiaomeng Lu, Hong Luo, Xuan Luo, Adrien Matray, Ben Matthies, Simon Mayer, Paolina Medina, Brian Melzer, John Mondragon, Christopher Palmer, Jonathan Payne, Cameron Peng, Lin Peng, Nagpurnanand Prabhala, Amiyatosh Purnanandam, Uday Rajan, Michael Reher, Jonathan Reuter, David Schoenherr, Scott Schuh, André F. Silva, Huan Tang, Lulu Wang, Shang-Jin Wei, Barry Williams, Emily Williams, Mungo Wilson, Carlo Wix, Yuhai Xuan, Liyan Yang, Constantine Yannelis, Motohiro Yogo, Yao Zeng, Shengxing Zhang, Yinuo Zhang, and Xiang Zheng, as well as participants at Princeton, Luohan Academy, PKU Econ, SUFE, PKU Guanghua, Bank of Canada, JHU Carey, TAMU, MSU, Oxford Saïd, UCI, Michigan Ross, Baruch, McGill, PHBS, UCL, Harvard, IMF, CB&DC JMC Workshop, DC&F Workshop, Inter-Finance PhD Seminar, NBER Chinese Economy Meeting 2022, MIT GCFP 2022, UConn Finance 2022, EFA 2022, NFA 2022, MFA 2022, EasternFA 2022, FIRS 2022, EoP XI 2022, OCC FinTech 2022, EBA PRW 2022, AMES 2022, AEDC 2022, AsianFA 2022, CFRC 2022, BFWG 2022, SEHO 2022, EFT 2022, FMCG 2022, ESDWS 2021, GAIF 2021, NZFM 2021, ICEA Inequality 2021, and SEL Fintech 2021. Thanks to Luohan Academy for a dissertation fellowship and access to the data, and to PIIRS for a research grant. Any errors are my own.

*Department of Economics and Bendheim Center for Finance, Princeton University, Princeton NJ 08542. Email: souyang@princeton.edu.

> *Digital payments also generate real-time data on sellers' businesses, the timing of cash flows, and buyers' purchasing habits, allowing payment providers to offer credit, savings, wealth management, collections, insurance, and other financial services. Where credit was once the way to draw in customers and offer a panoply of financial services, payments may be a safer channel for such upselling.*
>
> —Raghuram G. Rajan (2021). *All Eyes on Digital Payments.*

It has always been hard to provide financial services to the underprivileged, especially in terms of extending credit access.[1] Recently, there has been a global boom in credit provided by BigTech firms,[2] supplanting FinTech credit (Cornelli et al., 2021) and potentially reaching consumers who are unbanked or underbanked. At the same time, cashless payments—especially mobile payments provided by BigTech firms—effectively combine important technological advancements and have facilitated the transition to a cashless society worldwide.[3] These drastic changes give rise to two open questions of concern in academia and policy (Berg et al., 2021): Is the information from payment flows a causal factor behind the expansion of credit extended by BigTech? Does this expansion benefit consumers who were previously underserved by traditional financial institutions?

Whereas theory yields ambiguous predictions regarding the effects of payment adoption on consumer credit, my data reveal an answer. Conceptually, when a lender has more precise information on borrowers, the credit provided to the underprivileged can go up or down. On the one hand, the information can help them build a credit profile, which opens the door to financial services. On the other hand, the information can potentially render them more vulnerable to discrimination. I leverage a unique setting in China and use rich administrative data from Alipay, a leading BigTech platform that acts as a one-stop-shop for payment, credit, and many other financial

[1] The public and private sectors have proposed various solutions based on novel mechanisms or new technologies to solve this important issue. As perhaps the most notable example, the microcredit movement has achieved huge impacts but also faces limitations in scalability, cost reduction, and sustainability (Helms et al., 2006).

[2] The dominant and largest companies in the information technology (IT) industry such as Alibaba, Alphabet (Google), Amazon, Meta (Facebook), and Tencent.

[3] New technologies, including better collection and use of rich data (Agarwal et al., 2021a; Berg et al., 2020); more advanced credit risk models (Fuster et al., 2019, 2022); and financial accounts that are more accessible (Ouma et al., 2017) partially mitigate the limitations of traditional microcredit programs. Mobile payments have advantages in both data, models, and accessibility. First, payment records are by-products of daily purchases, which are rich, high frequency, and manipulation-proof. Second, cashless payment providers not only employ the most advanced machine learning and artificial intelligence technologies but also have access to data that facilitate model training and fully empower predictive credit risk models. Third, mobile phones have been widely adopted globally, which lowers the adoption cost of mobile payment and renders it almost universally accessible.

services for more than 1 billion active users.[4] Using a novel instrumental variable (IV) strategy, I show that cashless payment flow has a sizable impact on credit provision on both the extensive and intensive margins, especially for the previously financially underserved.[5] Distinct from the traditional credit card business model, which relies on credit repayment histories, the BigTech lender effectively employs the information on a borrower's creditworthiness revealed by the payment flow.

To quantify the value of this payment information, I construct and estimate a simple model of consumer finance with a BigTech lender. I simulate a counterfactual in which the lender only observes basic personal characteristics. I find that the availability of additional payment data significantly increases the average credit line provided by the BigTech lender to the borrowers. After the credit expansion, the default rate rises, and consumer surplus and lender profit increase. The structural analysis corroborates my IV analysis: Cashless payment flow facilitates credit provision, especially to the financially underserved.

Establishing a causal relationship between cashless payment and BigTech credit provision is challenging. First, it requires an exogenous shock to cashless payment activity, which is hard to identify, especially in mature markets. Second, I need detailed individual-level data on payments, credit, and investments, as well as information on the individual's sociodemographic conditions. Third, in order to focus on the credit supply, I will have to exclude credit demand factors from the observed credit line.

I address the first challenge by leveraging a natural experiment that provides exogenous variations of consumers' in-person Alipay payment, which is the staggered placement of Alipay-bundled shared bikes across different Chinese cities. I use bike placement as an instrument. The usage of shared bikes nudges users to make more in-person cashless payments with Alipay since both services require that consumers use Alipay to perform the same scanning procedure.

To address the second challenge, I base my analysis on administrative data from Alipay, which cover a representative sample and contain detailed information on users' personal characteristics and daily activities—consumption, credit access and usage, investments, shared-bike usage, and other relevant digital footprints. Linked household behaviors are measured at monthly frequency and recorded as individual-level panel data. A feature of the Huabei credit line enables me to address the third challenge;

[4]The *Huabei* credit line, a virtual credit card product provided by Alipay, is the largest consumer finance product in China as of 2020. It is the credit product I will focus on in this study.

[5]This cannot be directly achieved by performing a predictive exercise using historical data or conducting a field experiment. The former suffers from the manipulation critique of Bjorkegren et al. (2020), since consumers may strategically alter their behaviors when credit provision rules are encoded. The latter might have limited external validity, especially when it involves only a small population and runs for a short period.

different from a traditional credit card, it requires no active application and consumers instantly learn their qualification status and approximate credit line. This feature allows me to identify the credit provision effect from the supply side, which is immune to endogenous credit application motives from the demand side (Han et al., 2009; Brown et al., 2011).

I perform multiple tests to show the validity of the staggered placement of Alipay-bundled shared bikes in different cities as the instrumental variable. I show supporting evidence that this approach is likely to satisfy both the relevance condition and the exclusion restriction. The relevance condition requires a strong first-stage relationship between city-level bike placement and the in-person payment flow of Alipay users living in the city. My results confirm this view. The exclusion restriction condition requires that bike placement affect credit provision only through in-person cashless payment. I design tests to rule out potential concerns, including city-level common factors correlated with bike placement and credit provision, the selection bias of bike users, and the direct credit-revealing effects of bike usage.

The IV analysis section presents the empirical findings with respect to three parts of the study. First, I show that the exogenous increase in a consumer's in-person payment flow leads to more digital credit provided by Alipay and more credit take-up by the consumer. On the extensive margin, the use of in-person payment in a month leads to a 56.3% increase in the probability of getting credit access in the same month. On the intensive margin, for those with credit access, a 1% increase in the in-person cashless payment flow results in a 0.41% increase in the credit line. Given the exponential growth of the digital payment market in China, the accompanying credit expansion is enormous. The learning-by-doing story and the credit supply mechanism predict that consumers will change their borrowing behaviors. I find that more in-person payment flow leads to more credit take-up in both in-person and online settings. Although a more relaxed credit constraint can render a rational borrower weakly better off, it might lead to overspending for consumers with behavioral biases, such as self-control and forecasting problems (Ausubel, 1991; Melzer, 2011; Di Maggio and Yao, 2020). With an analysis of detailed transaction category data and an event study that evaluates the effect of credit access on spending, I do not find significant evidence of compulsive spending or overspending.

The second part of the IV analysis investigates the channels through which the in-person cashless payment flow facilitates credit provision. I find evidence that the transaction data contain valuable information for credit evaluation, which confirms the long-standing discussion in the banking literature that highlights the bank's cost advantage in making loans to depositors based on information in their deposit history (Black, 1975; Fama, 1985). Even when we focus on noncredit payment flow, which does not contain information on credit usage and repayment, we still find a significant and positive effect on credit facilitation. At first sight, this can be surprising, since banks usually rely on the credit usage and repayment history the credit bureau provides to make decisions on consumer credit provision. On reflection, however, it makes sense, especially regarding the recent theoretical discussion of open banking (Parlour et al., 2022; He et al., 2022) and recent empirical evidence that demonstrates the potential of payment footprints over traditional credit scores in credit evaluation Rishabh (2022). Cashless payment flow is associated with both information on the consumer's creditworthiness and the enforcement power of the BigTech lender (Brunnermeier and Payne, 2022). It is hard to fully separate these two forces. I use the consumer's assets under management on Alipay as a proxy for collateral, since Alipay can potentially freeze the account if the user does not repay on time. I find that the payment flow information channel still holds when I control for this collateral proxy.

In the third part of the IV analysis results, I investigate the implications of digital payment for financial inclusion. Using an illustrative theoretical example, I show that digital payment adoption, as an information shock to the lender, can potentially lead to opposite credit access outcomes for borrowers with lower creditworthiness. Less creditworthy borrowers can get a higher credit line or lower credit line after the information shock, depending on the parameter values. I find that the financially underserved gain more access to credit after adopting in-person cashless payment. There is a traditional view in China that the less educated and the older tend to be financially underserved.[6] My data confirm this view: These groups have fewer financial activities and lower financial literacy. The exogenous increase in the in-person cashless payment flow results in an increase in credit provision mainly to the less educated and older segments of the population.

In the model-based analysis section, I estimate a simple model to evaluate the real effects of the digital payment adoption shock and quantify the value of using payment data in the consumer credit market. Comparing the case in which the lender knows the consumption information in addition to personal characteristics with the counterfactual case in which the lender knows only borrowers' characteristics, the availability of payment data increases the credit line by 57.7% on average. Although the default rate increases by 13.3%, the annual consumer surplus is improved by 151.2 CNY per capita—which is 0.5% of the median per capita disposable income of residents in China—and the annual lender profit expands by 62.4 CNY per capita. Consistent with the IV analysis, the less educated and the older benefit more in the process.

[6]See the Bank of Finland Institute for Emerging Economies (BOFIT) article, "Chinese Increasingly Likely to Attend University, Nearly All Young Adults are Literate" (May 21, 2021).

An emerging literature examines the relationship between digital payment and digital credit. Berg et al. (2021) provide an extensive review of FinTech lending and highlight open questions regarding the role of payment data in the credit market. A growing literature on BigTech lending addresses these questions with a focus on business loans.[7] To my knowledge, this is the first paper to empirically reveal the causal effects of payment flow information on facilitating consumer credit provision.[8] A recent theoretical paper by Parlour et al. (2022) studies a model of competition between financial intermediations for payment processing, in which the premise of the analysis is that payment flow data contain information on the consumer's credit quality. My paper provides evidence that directly supports Parlour et al. (2022)'s assumption regarding the informativeness of each consumer's payment flow. Another closely linked article is by Ghosh et al. (2022), who uncover the synergy between FinTech small-business lending and cashless payments using theoretical and empirical analyses. I show that consumers are less strategic than firms when deciding whether to adopt cashless payment: Even a tiny nudge toward digital service usage can lead to a significant shift in the long-run choice of payment instruments, and the difference in setups results in opposite predictions. Ghosh et al. (2022)'s theory suggests that better firms benefit more from cashless payment adoption due to the information-revealing effect, whereas my paper shows that it is the financially underserved who get more credit access after the adoption of cashless payment.

My paper contributes to the literature on the effects of payment technology adoption on consumers. Digital payment products, including debit cards and mobile payments, can reduce transaction costs, monitoring costs, and travel costs, which leads to further changes in consumer banking (Mbiti and Weil, 2015); household savings (Bachas et al., 2021); risk-sharing (Jack and Suri, 2014; Riley, 2018); risk-taking (Hong et al., 2020); consumption (Suri and Jack, 2016); corruption (Tang et al., 2022); and crime-related risk (Economides and Jeziorski, 2017). So far, this literature has largely focused on the

[7] Frost et al. (2019) and Cornelli et al. (2020, 2021) discuss recent trends in BigTech finance and demonstrate BigTech lenders' advantage in using non-traditional information and serving unbanked borrowers, especially in markets with less competitive banking sectors. Liu et al. (2022) show that a BigTech lender mainly serves borrowers' short-term liquidity needs and facilitates credit to underbanked businesses without experiencing more severe adverse selection. Beck et al. (2022) and Gambacorta et al. (2022) use business lending data from the Ant Group to show that business-related transaction data can act as a substitute for collateral and facilitate firms' credit access. Hau et al. (2019, 2021) and Chen et al. (2021c) study BigTech credit to e-commerce firms using data from the Ant Group and highlight the positive effects of BigTech lending, including the mitigation of financial frictions, boost in sales growth, and reduction of sales volatility.

[8] The consumer credit studied in this paper differs from business loans in several respects. First, it is a consumer credit card used by consumers for direct purchases. In contrast, business loans can be used for investments or transfers, since businesses can directly obtain cash equivalents. Second, the size of consumer credit is much smaller than that of business loans. For example, the average consumer credit line is 6,259 CNY in my sample, while the average credit limit of the uncollateralized BigTech business loan is 71,963 CNY.

cost-reduction effects of new payment products and rarely on the value of the payment data accumulated as a result of digitalization. My paper addresses this point by quantitatively assessing the real effects of payment data usage in the consumer credit market. My paper is also among the first to analyze a BigTech payment app, which provides not only payment services but also a large set of data-based financial services and daily-life services. It thus provides suggestive evidence that shows the potential power of service bundling and interoperability, which naturally arises for digital financial platforms (Brunnermeier and Payne, 2022; Benetton et al., 2022; Agarwal et al., 2021b).

This paper also adds to the literature on the real consequences of consumer credit. Although there is an increasing variety of consumer finance products, there is less empirical consensus on the effects of these products on consumers' financial health and welfare (Zinman, 2015). On the one hand, if borrowers are rational, revealed preference logic suggests that they would take advantage of consumer credit to relax their financial constraints and smooth their consumption; they should be at least weakly better off with their expanding financing choices. Empirical studies find positive effects of consumer credit even when the credit is expensive (Karlan and Zinman, 2010; Morse, 2011). The large literature on financial inclusion supports the view that better access to financial services is especially beneficial for disadvantaged groups.[9] On the other hand, if borrowers have behavioral biases, they might overspend, overborrow, and face negative financial consequences (Medina, 2021). Time-inconsistent preferences (Laibson et al., 2000); optimal expectations (Brunnermeier and Parker, 2005); and limited attention (Stango and Zinman, 2014) can potentially lead to high consumer debt burdens. Some studies find evidence of the negative effects of consumer credit in the payday loan market (Melzer, 2011), credit card market (Ausubel, 1991), mortgage market (Begley and Purnanandam, 2021), FinTech lending market (Di Maggio and Yao, 2020), and "Buy Now Pay Later" (BNPL) market (deHaan et al., 2022). Despite the high popularity and growth of BigTech consumer credit, archival research on it is scarce. My study fills this gap with transaction-level data and shows that more digital payment activities and more BigTech credit access do not significantly affect consumers' compulsive spending. Consumers increase their spending levels after getting BigTech credit access without a reversal. This is consistent with the "flypaper effect" documented by Di Maggio et al. (2022) in the BNPL market. My structural estimation indicates that BigTech credit expansion leads to higher consumer welfare and lender profit, with a cost of higher default rates.

Several studies examine the determinants of consumer credit. Whereas collateral has always been important for the mortgage market (Mian and Sufi, 2011) and business

[9]See Brown et al. (2019); Célerier and Matray (2019); Chopra et al. (2017); Doornik et al. (2021); Erel and Liebersohn (2022); Stein and Yannelis (2020); Di Maggio et al. (2021) for extensive discussions on this point.

lending (Kiyotaki and Moore, 1997; Rampini and Viswanathan, 2020), many consumer credit products are unsecured. Credit usage and repayment history information from the credit bureau is critical for credit evaluation (Chatterjee et al., 2020; Ordoñez et al., 2019) and has traditionally driven the household finance market. As technology advances, more hard information becomes available and facilitates new business models (Liberti and Petersen, 2019).[10] An emerging literature uses predictive exercises to demonstrate the wide-ranging potential of digital footprints (Berg et al., 2020; Agarwal et al., 2021a; Rishabh, 2022) and machine learning models (Di Maggio et al., 2021; Fuster et al., 2022) in credit evaluation. My paper complements these studies by demonstrating the value of payment data for credit provision in the real business environment. I show that this effect of payment flow information goes beyond credit usage, repayment, and the enforcement power from the assets under management on the BigTech platform. By estimating a simple model and performing the counterfactual analysis, I show that the availability of payment data can potentially increase consumer welfare and lender profit simultaneously.[11]

The paper is organized as follows. Section 1 provides some institutional background on the Alipay platform and the dockless bike-sharing industry in China. Section 2 describes the data and the identification challenge and provides evidence on the validity of the instrumental variable. IV analysis results are in Section 3, in which I analyze the relationship between cashless payment flow, credit provision, and financial inclusion. The model-based analysis is in Section 4, in which I construct a simple model, estimate the key parameters, and study the information value of payment data. I conclude in Section 5.

1 Institutional Background

Two observations motivate this study: the fast development of cashless payment systems worldwide and the rise of consumer lending by FinTech and BigTech companies in many countries. First, cashless payment systems—especially in-person mobile payments—have achieved remarkable success in less than a decade. China is a typical example. Figure 1 highlights China's dominance in mobile payment in 2019 and 2021, with the highest user penetration rate and income-adjusted annual transaction value

[10] Liu et al. (2022); Tang (2019); Buchak et al. (2018); Beaumont et al. (2022); Howell et al. (2021); Huang (2022) discuss the comparative advantages of FinTech and BigTech lending models relative to traditional banks. Regulatory arbitrage, alternative data access, big data technology, liquidity, process automation, and convenience play important roles.

[11] Thus, my paper is also related to discussion of the implications of open banking (He et al., 2022; Goldstein et al., 2022; Babina et al., 2022); the welfare impacts of information frictions in credit markets (DeFusco et al., 2021; Jansen et al., 2022); and data sharing and the value of data (Chen et al., 2021a,b)).

per user.[12] However, other countries also increased their mobile payment adoption significantly in two years. Figure (a) shows China's clear lead in 2019, while Figure (b) shows a narrower gap in 2021. The UK and the US had almost reached China's level, and developing countries such as Nigeria and India had also improved their performance. The changes were largely driven by mobile payments for in-person transactions. China was massively shifting from a cash economy to a cashless economy. Figure 2 shows that from 2012 to 2018, the annual transaction volume of China's mobile payment system increased from 4% of GDP to 302%, while the corresponding measure of US card payments stayed below 34% of GDP. China's mobile payment market provides a unique setting to study the impact of cashless payment and has great implications for other countries and the future.[13] Second, at the same time, BigTech credit is surpassing FinTech credit globally (Cornelli et al., 2020). China's experience is crucial to understand, as it has the largest market for both types of credit, dominated by Alipay. In a representative sample of Alipay users, I find that 72% have access to a Huabei credit line, of which more than 95% have used it at least once and have an average monthly credit usage of 533 CNY.[14] The credit product is quite inclusive, even among users who do not have a bank-issued credit card on file; of those, 64% have Huabei credit line access.

China is rapidly becoming the world leader in mobile payment, and the path it takes is quite pioneering.[15] Unlike the mobile-phone-based payment system that relies on SMS text messages—e.g., M-PESA in Kenya—or card-complementing mobile payment systems, such as Apple Pay or Google Pay, China's mobile payment system is based on so-called "super apps." Among these, the most notable are Alipay and TenPay, which provide an all-in-one digital experience to users with both in-house services and integrated third-party services. This paper studies mobile payment in China by analyzing Alipay's proprietary data.

1.1 The Alipay Platform

Alipay is the largest digital payment services provider as measured by total payment volume in China, which reached RMB 118 trillion from July 2019 to June 2020. It is a

[12] Also see the World Economic Forum article by Katharina Buchholz, "China is Fast Becoming the World Leader in Mobile Payment" (May 15, 2019).

[13] There has been a global trend of going cashless in in-person payments, and the pandemic might further speed up the process. See the *Forbes* article by Len Covello, "How the Pandemic Made Contactless Payments the New Normal" (April 15, 2021).

[14] This is roughly 80 USD, about 20% of the median per capita disposable income of residents in China.

[15] Han and Wang (2021) construct a model to explain this leap-frogging phenomenon in mobile payment adoption across different countries. Cong and Mayer (2022) models the competition between fiat money and digital money with a game-theoretic framework. Ouyang and Peng (2022) analyze the drivers of the growth in mobile payment in China both empirically and theoretically.

third-party mobile and online payment platform launched by the Alibaba Group in China in 2004. As of late 2020, it had over 1 billion users, 80 million merchants, and more than 2,000 partner financial institutions for digital payment and digital financial services, including unsecured consumer credit. Alipay has always been the principal means by which buyers transact with sellers on Alibaba's platforms since its launch. It has grown explosively since 2016, in both the number of users and transaction volume.

China has switched from a cash economy to a cashless economy in less than one decade, during which Alipay has played an important role. Nowadays, consumers in China rarely carry cash. Instead, they use Alipay and TenPay to pay for almost everything, including taxis, bills, e-commerce purchases, and even purchases from small street vendors. Alipay has become a platform that enables merchants and consumers to complete transactions for almost all online and in-person payments. It also acts as a one-stop shop for digital payment, digital financial services, and a broad range of daily life services. Using Alipay, a consumer can access 1,000 daily life services and more than 2 million mini programs that provide mobility services, local services, municipal services, and many others, without having to download additional apps.

In Figure A.1, an illustration from the prospectus of the Ant Group lists the typical use cases available via the Alipay app. From the Ant Group's standpoint, the foundation of all such services is digital payment; the provision of other digital financial services, including consumer credit, wealth management, and insurance, are based on it. Consumers could fund payments for major uses through the e-wallet account balance, the Huabei credit line, and linked bank card accounts.

Here, Huabei is a virtual credit line product that can only be accessed via Alipay. It has some unique features that simplify and clarify the research environment. In this paper, I measure Alipay's credit provision based on access to the Huabei credit line. As of late 2020, it is the largest digital consumer credit product by credit balance in China. First, Alipay users can instantly find out if they qualify for Huabei and their approximate credit line, unlike traditional credit cards that require an application and a waiting period. This feature allows us to focus on the credit supply without worrying about the endogenous changes in the demand side that might occur for traditional credit card products. Second, there is almost no price discrimination. Most users pay a daily interest rate of 0.05%, or 18.25% annually. This helps us to focus on the quantity dimension of the credit instead of the price of the credit. Third, the risk management for the product was excellent, so that we can worry less about the possibility of unsustainable credit expansion. As of June 2019, Huabei's delinquency rate was 1.16%, while China's publicly listed banks' credit cards had rates between 1.21% and 2.49%. This is impressive, since we usually consider credit card users in

China to be more creditworthy. It suggests that Alipay might have used data effectively to improve credit evaluation.

Huabei has some other features and statistics. Once an Alipay user is granted access to Huabei, her credit line is instantly available at the point of sale. The whole process is fully automatic. The minimum credit line is as low as 20 CNY (roughly 3 USD), and Alipay offers consumers an interest-free period of up to 40 days after the corresponding purchase. Consumers have the option to pay in monthly installments over 3 to 12 months after the interest-free period. From July 2019 to June 2020, the average Huabei outstanding balance was around 2,000 CNY.

1.2 The Dockless Bike-sharing Market in China

Unlike traditional bike-sharing systems that offer rental bikes that are docked in stations, dockless bike-sharing platforms offer users more convenient services. They can use bike-sharing apps or mobile wallet apps to scan the QR code on the bike's smart lock and unlock an available shared bike in seconds. After finishing the trip in any authorized area, they can reset the lock easily; the bike is then available to other users.

The bike sharing service in China boasts two major features. Firstly, it is affordable, with a usage rate of 0.23 USD for the first 15 minutes and 0.08 USD for every subsequent 15 minutes, and an unlimited plan available for approximately 3 USD per month. Secondly, it has a massive user base, with 260 million shared bike users in late 2019 and over 400 million users for Hellobike alone in 2021.

The first dockless bike-sharing firm in China was *ofo*, which was founded in 2015 in Beijing. It started as a two-sided platform that enabled students to share their bikes and ride others' bikes on campus, and later shifted to a one-sided platform supplying GPS-tracked dockless bikes to users of its bike-sharing app (Cao et al., 2018).

Since late 2015, the bike-sharing industry in China has attracted investment from venture capital (VC) funds and BigTech firms, and has undergone exponential growth (Figure A.2). According to data from China's transport ministry, there were 23 million shared bikes from 77 companies in hundreds of Chinese cities as of early 2018, when ofo and Mobike accounted for 95% of the market in total.

There were rises and falls among bike-sharing service providers. Although ofo was the first player in the bike-sharing industry and at one time dominated it, it later incurred a large amount of debt it was unable to pay, and had stopped operating bike rentals as of 2020. In contrast, *Hellobike* was a small bike-sharing provider in 2017, but has since become the largest bike-sharing service provider in the world as measured by the number of total rides in 2020.

1.3 Digital Payment Competition and Dockless Bike-sharing Market

In 2013, non-cash retail payments in China amounted to less than RMB 50 trillion, and almost all were debit card or credit card transactions. At that time, in-person mobile payment services such as Alipay or TenPay accounted for only a tiny fraction of the non-cash retail transaction volume. The market size of in-person mobile payments grew gradually at first and took off quickly after 2016. As of 2019, non-cash retail payments in China exceeded RMB 350 trillion, with more than RMB 200 trillion attributable to in-person transactions made through mobile payment service providers.

Ant Group's Alipay and Tencent's TenPay are the two major digital payment service providers in China. As of June 2020, Alipay had about 55% and TenPay had about 40% of the market share by transaction volume. Alipay and TenPay have competed fiercely and invested heavily to grow their market size and share.

A strategic move by mobile wallets has been partnering with bike-sharing companies, since the digital payment system can provide infrastructure for bike-sharing services. The high-frequency usage of bike-sharing services, in turn, can encourage users to adopt mobile wallet use for other payments in daily life.

Because of the synergistic effect between digital payment and bike-sharing services, Alibaba and Ant Group invested more than 0.5 billion dollars in ofo and more than 3 billion dollars in Hellobike; ofo was once the largest player and Hellobike is now the largest player in the bike-sharing industry. In return, these bike-sharing services are deeply bundled with Alipay. By taking advantage of the mini-programs within the Alipay system, Alipay users can unlock shared bikes by scanning the QR code on the bike with Alipay directly, without downloading the specific bike-sharing app or manually entering personal information. This relationship is exclusive; a WeChat user is unable to directly unlock a shared bike operated by Hellobike using TenPay. What is more, for Alipay users who have a high enough credit score in Alipay's credit scoring system, the deposit for using shared bikes can be waived. According to the IPO prospectus of Hello Inc., the company that operates Hellobike, "The popularity of our service and our rapid business expansion, in turn, contribute to the prosperity of the ecosystem built upon such payment and digital infrastructure."

The bike-sharing and mobile payment markets boomed from 2016 to 2020. This offers a unique setting to study the causal effects of cashless payment, since Alipay-bundled shared bikes were placed in different Chinese cities at different times, causing exogenous adoption shocks to Alipay users in those cities.

2 Data and Identification

2.1 Data Description

It has always been challenging to obtain a suitable data set to study the relationship between payment flow and consumer lending. It requires granular data with linked payment and credit information. It is even harder to study it in a dynamic setting. I overcome these challenges by using proprietary panel data at individual and year-month level from Ant Group, which contain detailed information on not only broad payment and credit activities, but also rich personal characteristics.

The main dataset used in the study consists of panel data that include 41,485 randomly selected Alipay users who have at least one in-person transaction in the sample period of May 2017 to September 2020. For each user, I observe both the static characteristics of gender, education, year of birth, and so on, and time-varying measures, such as in-person payment flow, online payment flow, bike-riding activity, credit provision, and credit usage. Another important dataset used in the study consists of city-level panel data on the placement of Alipay-bundled shared bikes.

Table 1 reports a summary of the distribution of the sample in multiple dimensions. The first set of characteristics is at the individual level and covers 41,485 Alipay users. The average user in the sample was born in 1983, with payment activities in 32 months of the 41 months from May 2017 to September 2020. Roughly 54% of users in the sample are male. About 88% of sampled users do not have a bachelor's or higher degree, and 29% of users used Alipay-bundled shared bikes at least once during the sample period. The second set of measures is at city and year-month level. In the average sampled month, the average city has a log transformed number of placed shared bikes of 7.08. The third set of variables is at individual and year-month level. In the average sampled month, the average user has a 62% probability of having access to Alipay's virtual credit card, a log-transformed credit line of 7.88, a log-transformed in-person Alipay payment amount of 5.70, and a log-transformed online Alipay payment amount of 5.76, where the credit line and payment flows are measured in CNY. For the average user in the average month, 34% of in-person Alipay payments and 33% of online Alipay payments are paid using the virtual credit card, and 3% of in-person Alipay payments and 1% of online Alipay payments are for compulsive spending, including cigarettes, games, lotteries, or live streaming services.

Alipay was widely adopted and intensively used by consumers at the end of my sample period. As of September 2020, the median size of a sampled user's assets under management in the Alipay wallet was 8 CNY, and the average size was 5,521 CNY. The median is small since many users do not keep a positive balance on their Alipay accounts and use linked bank cards to make purchases instead. In my data, the median monthly Alipay transaction amount was 238 CNY at the end of the sample period, and the average was 2,628 CNY. As a comparison, the average per capita monthly disposable income in China was 2,682 CNY in 2020.

2.2 Identification Challenge

Several endogeneity issues arise in addressing the causal relationship between cashless payment and credit provision. For example, simultaneity can occur when there is synergy between the adoption of cashless payment and credit provision by the payment service provider (Ghosh et al., 2022) or other factors that may simultaneously affect payment and credit. Omitted variables that potentially bias the estimates can also be present. Section A.1 provides an econometric framework to illustrate the economic environment and the endogeneity issues.

Exogenous variations in digital payment adoption can help address these issues. However, they are in general hard to identify, especially in countries with developed financial systems and widely adopted digital payments. For example, debit and credit cards are already quite popular and accessible in the US, and thus cashless payment activity is endogenously determined; those who use cards for daily purchases are notably different in nature from those who use cash. In contrast, mobile payment is quickly being adopted in China and provides a unique setting to generate exogenous variations in cashless payment adoption across different cities over time. I explain how I address endogeneity issues using an instrumental variable approach in the following sections.

2.3 The Nudge Effect

This subsection provides some direct empirical evidence that supports the story illustrated in Figure 3. I use the placement of Alipay-bundled shared bikes across cities as a novel instrumental variable to alleviate endogeneity concerns. Although Alipay, the mobile payment leader in China, grew rapidly in recent years, there were also staggered placements of Alipay-bundled shared bikes across different cities; this led to exogenous shocks to bike users' adoption of Alipay. When there are more Alipay-bundled shared bikes placed in the city, the bike-sharing service becomes more valuable for bike users, which in turn motivates them to use Alipay more frequently to unlock bikes by scanning the QR code on the bike. This frequent usage of Alipay nudges users to trust Alipay and be comfortable using it not only for bike-related spending but also for other in-person payments. After all, scanning the QR code on a shared bike to unlock it and scanning the QR code of a merchant to make a payment are the same in terms of the procedure.

I provide direct evidence for the logic flow illustrated above, which can be used as a sanity check. First, I show that when more Alipay-bundled shared bikes are placed in a city, individuals living in that city have higher bike-riding activity. Second, I show that after an individual adopts the use of shared bikes, her in-person payment flows that are unrelated to shared bikes are likely to increase abruptly.

Table A.1 presents ordinary least squares (OLS) estimates from regressions that focus on the sample of Alipay users who have used a shared bike at least once in the sample period, and Columns (2) and (3) focus on the months in which bike users use bikes. The results show the positive relationship between the city-level placement of shared bikes and individual-level usage of shared bikes on both the extensive margin and intensive margin. The estimates suggest that on the extensive margin, for sampled bike-riding Alipay users living in city c, having a 1% increase in city-level bike placement in city c in month t increases a user's probability of using shared bikes by 0.028%. On the intensive margin, for bike users in the months they use bikes, the 1% increase in bike placement in month t leads to an increase in the bike user's number of bike rides of 0.082% and an increase in her total distance for bike rides of 0.120% in month t. When more bikes are placed in a city, finding an available shared bike becomes easier for bike users, and they are expected to have higher bike-riding activity. In addition, as Cao et al. (2018) demonstrate, since the dockless bike-sharing system is a one-sided network with positive network effects, there might also exist indirect effects, whereby more bike-riding activity by one user also increases others' bike-riding activity. Both the direct and indirect channels lead to a positive relationship between the city's bike placement and the bike-riding activity of bike users living in the city.

Next, I provide evidence on the nudge effect of shared-bike adoption on in-person payment activity. Table A.2 shows the strong correlation between bike usage and in-person cashless payment flow with regressions. This does not evolve in a gradual manner. Figure 4 is a graphical illustration of the effects of bike adoption on in-person payment flow that is unrelated to the use of Alipay-bundled shared bikes. It uses an event study framework, in which the event for individual i is her bike adoption and t corresponds to the number of months after the individual's month of the first use of Alipay-bundled shared bikes. The reference time 0 indicates the end of the month of each user's bike adoption. The figure plots the β_τ coefficients estimated in the regression:

$$log(1 + \text{In-Person Non-Bike Payment Flow})_{i,t} = \alpha_0 + \sum_{\tau=-5}^{4} \beta_\tau \cdot \mathbb{1}(t = \tau) \cdot \mathbb{1}(\tau \neq -1) + \beta_5 \cdot \mathbb{1}(t \geq 5) + \delta_i + \mu_t + \varepsilon_{i,t} \quad (1)$$

where δ_i represents city fixed effects and μ_t represents year-month fixed effects. For each bike user, the sample only covers periods in which the event time t is not earlier than -5. Compared with the benchmark month, in-person non-bike payment flow increases by more than 80% in the month of shared-bike adoption and stays at a level more than 30% above the benchmark level in the following months. Although the bike-adoption decision itself is endogenous, this sharp contrast of in-person non-bike payment flow before and after shared-bike adoption suggests that it is the use of Alipay-bundled shared bikes that leads to a shift in payment habits. Otherwise, the change should not be so abrupt around the heterogeneous shared-bike adoption date of users, especially with respect to individual and year-month fixed effects. This phenomenon is likely to be caused by switching from paying with cash or other payment instruments to paying with Alipay, rather than by sharply changing the level of consumption after shared-bike adoption. Note that this is not mechanically driven by the people who register for Alipay only to gain access to Alipay-bundled shared bikes. As Figure A.3 shows, the vast majority of sampled Alipay users either adopt Alipay-bundled shared bikes after being an Alipay user for more than 1 year or do not use the shared bikes at all in the sample period, and only 1% of users start to use Alipay-bundled shared bikes in their first year of Alipay usage. Thus, the mechanical effect should be negligible.

2.4 Validity of the Instrumental Variable

In this subsection, I provide empirical evidence that supports the use of city-level bike placement as a valid instrument for individual-level in-person cashless payment that is likely to satisfy both the relevance condition and the exclusion restriction condition. First, I find a strong relationship between bike placement in a city and the in-person cashless payment flow of Alipay users living there. Second, I show that bike placement is likely to affect Alipay credit provision only through in-person cashless payment.

2.4.1 The Relevance Condition

There are concerns that city-level bike placement might not be a strong instrument for individual-level in-person cashless payment flow, especially when granular controls are added. The data show that this relevance condition can be robustly satisfied, and the results suggest that bike placement acts as an exogenous shock to Alipay users' in-person payment through the nudge effect described in the preliminary analysis.

Panel A of Table 2 shows the effects of city-level placement of shared bikes on individual-level in-person payment flow. Column (1) shows that when the bike placement of city c in month t increases by 1%, the in-person payment flow of the individuals living in the city increases by 0.039% on average. The relationship is quite strong, even when both individual and year-month fixed effects are controlled for and when standard errors are double clustered by city and year-month levels. Individual fixed effects can capture the time-invariant determinants of in-person payment activities for everyone—such as financial literacy, digital literacy, and wealth level—while year-month fixed effects can capture the time-varying determinants of in-person payment activity, such as workday effects and holiday effects.

A closer look in column (2) reveals that this positive relationship between bike placement and in-person payment flow only exists for bike users, but not for users who have never used Alipay-bundled shared bikes. This result can be regarded as a placebo test that supports the view that it is bike placement that affects Alipay users' in-person payment through bike usage. And it makes sense that for non-bike users—especially those who do not know how to ride a bike—regardless of how many shared bikes are placed around them, their payment activities should not be directly impacted. This test also helps rule out stories whereby the positive relationship is driven by some unobserved common factors that affect the whole population in the local area—e.g., local growth potential or local infrastructure plans—and that simultaneously correlate with the city's bike placement and city residents' in-person payment flow.

Column (3) focuses on bike users and shows the results of the regression with a specification that further adds city times year-month fixed effects, which remove all unobserved time-varying heterogeneity across cities, such as differences in local business cycles, different levels of local Alipay penetration, different local trends in bike placement, or aggregate variations that could arise from the placement of shared bikes. Identification of the coefficient relies on comparing the in-person payment flow of bike users in response to bike placement with that of a control group of non-bike users within the same city, with the static characteristics of the individuals controlled for at the same time.

The intensive margin analysis supports the mechanism whereby bike placement exogenously affects in-person payment flow through bike usage. Differences in the response of in-person payment flow to bike placement not only exist between bike users and non-bike users, but also before and after adoption of Alipay-bundled shared bikes within the same bike user. Without the variation in bike placement, the bike-adoption decision itself does not have a significant effect on the in-person payment flow, which alleviates concern about selection issues in the endogenous timing of bike adoption. For bike users, after the bike adoption, a 1% increase in bike placement results in a 0.051% increase in the in-person payment flow.

2.4.2 The Exclusion Restriction Condition

The identifying assumption is that bike placement affects digital credit provision only through in-person cashless payment. Three major concerns arise regarding satisfaction of the exclusion restriction by using the bike placement instrument. The first concern is that there exist factors that correlate with bike placement and credit provision at the same time. The second concern is that the use of Alipay-bundled shared bikes can directly reveal the creditworthiness of consumers and affects Alipay's credit provision. The third concern is that bike placement is predictable or clustered within a short time, which renders it not as exogenous as required.

The first concern is the existence of common factors that are correlated with bike placement and credit provision at the same time. For example, some time-varying growth potential for a city could attract the attention of both bike-sharing companies and Alipay; as a result, the likelihood of bike placement and the level of credit provision would increase at the same time. Panel B of Table 2 provides reduced-form results on the influence of bike placement on credit provision, and indicates that the positive relationship between bike placement and credit provision is unlikely to be driven by common factors unrelated to the bike-riding channel. Column (1) shows that the higher the bike placement shock in a city, the higher the credit line the individuals living in the city receive. In this setting, individual fixed effects and year-month fixed effects remove static heterogeneity across individuals and time-varying macroeconomic variations.

I further separate Alipay users into bike users and non-bike users and explore the heterogeneous effects of bike placement on their digital credit lines in column (2). It shows that the reduced-form positive effect of bike placement on credit provision only exists for bike users. The fact that bike placement has a positive effect on one group but not on the other group can seem surprising, especially when the difference between the two groups is quite small. In the definition used here, the only difference between a bike user and a non-bike user is whether the person has used Alipay-bundled shared bikes at least once during the whole sample period. The suggested mechanism explains the phenomenon very well, whereby bike placement first leads to more bike usage, then increases in-person payment flow, and finally results in more credit lines. It also helps reject the story whereby some factors correlate with both bike placement and credit provision, since the usual common factors are unlikely to affect bike users and non-bike users in different ways—especially when it is extremely inexpensive for an Alipay user to be a bike user as defined herein.

Column (3) focuses on bike users and reports results of the regression with individual fixed effects and city times year-month fixed effects. Although the timing of bike adoption is endogenous, the dummy variable that indicates whether the bike user has adopted shared bikes does not imply a higher credit line, which suggests that the timing itself does not play an important role in credit provision. The interaction term of the dummy variable and the bike placement, however, has a significant positive effect on the credit line, and this is consistent with the bike usage channel documented above.

Although the cost to become a bike user is low, one could argue that bike users and non-bike users have very different characteristics, and it is these associated characteristics, instead of bike usage itself, that lead to the difference in the reduced-form effect of bike placement on credit provision. To rule out this channel, I first screen the personal characteristics that are strongly associated with the bike-user classification, then check the heterogeneous effects of bike placement on credit provision along these dimensions. Table A.3 shows the regression results on the relationship between personal characteristics and the choice to become a bike user. Across different specifications, several personal characteristics are indeed correlated with the bike user dummy, including education, age, Alipay experience, gender, and indicators for whether paying with the user's real name or whether using their own account. Table A.4 reports the heterogeneous effects of bike placement on in-person payment flow and credit provision. The bike placement variable interacts with both the bike user dummy and the measure of personal characteristics selected from Table A.3. Panel A reports the OLS regression results in which the dependent variable is $log(1 + \text{In-Person Payment Flow})_{i,t}$, while Panel B shows the corresponding results in which the dependent variable is $log(1 + \text{Credit Line})_{i,t}$. Each column uses a different personal characteristic measure. Even though personal characteristics such as education, age, and gender all seem to be much harder to change than the status of being a bike user, across all specifications, the heterogeneity mostly arises from the dimension of the bike-user dummy. These results suggest that it is the bike-usage-associated behaviors, instead of the selection issue, that matters most in the effects of bike placement on in-person payment flow and credit provision. It is unlikely that bike users are a special group of individuals who benefit from the shock to the Alipay credit line simply because they have different personal characteristics, especially when everyone can easily join this group.

Among the common factors, there is a concern about the impact of the bike-placement shock on the local economy. Since dockless shared bikes offer great convenience to users and the number of bike users is large, some might worry that bike placement brings new business opportunities and affects the local economy or fiscal policy, which further leads to an increase in credit provision. Table A.5 shows the relationships between bike placement and the variables associated with local economic conditions. Under city fixed effects and year-month fixed effects, the coefficients for all specifications are small and insignificant, which indicates that bike placement is unlikely to have macroeconomic impacts.

The second concern is the direct revelation of creditworthiness by bike usage. Some institutional backgrounds and facts help alleviate this concern. First, Alipay is only a strategic partner with the bike-sharing companies and is unlikely to use third-party data directly as the model input. The bundling also seems to be limited, since the official bike apps support multiple mobile wallets and Alipay is not required for bike usage. Second, the cost of bike usage is very low, which renders the activity easy to manipulate. If the direct effect on credit provision is large and there exist some manipulations, the Alipay company, which is very sophisticated and advanced in technology, will fix those issues in equilibrium. The average cost of bike usage is as low as 0.23 USD for the first 15 minutes and 0.08 USD per 15 minutes after that. The monthly unlimited plan is only 3 USD, which can be regarded as an upper bound for the monthly bike spending of a rational user. Third, the user base is quite large, given that bike users are unlikely to be very selective. The size of the shared-bike user base in China is as large as 260 million as of late 2019, and Hellobike claimed to have over 400 million registered users as of 2021.

Table 3 furthers shows that bike usage is more like a nudge for payment activity and credit line, instead of a proof of creditworthiness. I separate bike users into two categories: the one-time bike user who has used Alipay-bundled shared bikes only once during the whole sample period, and the repeat bike user who has used the bikes at least twice in the data. Even if bike usage itself reveals some information about creditworthiness in the long run, using the bike once should not be very informative. Columns (1) and (3) show that bike placement has no significant effect on in-person payment flow and the credit line of the non-bike users, but has strong positive effects on in-person payment and credit line of one-time bike users, even though the difference between these two groups is only one bike-riding activity. Moreover, although the effects are stronger for repeat bike users, the difference in the effects between one-time bike users and repeat bike users is relatively small. Columns (2) and (4) indicate that the patterns are very robust, even when city times year-month fixed effects are added to the specification.

The third concern is about the bike placement process. If it is a predictable process or is clustered within a short period for all cities, it is more likely that it will correlate with other factors that are associated with credit provision. From the perspective of bike-sharing companies, it is more beneficial for them to make bike placement a staggered and unpredictable process, and the empirical evidence supports this. There is anecdotal evidence that what bike-sharing companies care most about is local competition and their own operational efficiency, and this could lead to heterogeneous overall strategies. For example, bike-sharing companies such as Mobike and ofo focused mostly on big cities in the beginning and gradually expanded to smaller cities, while Hellobike started bike placement in small cities first to avoid competition and then gradually expanded to larger cities. Regardless of which cities they decide to target first, bike-sharing companies always have an incentive to quickly place their shared bikes in the local market because it helps them build local market power and avoid competitors who may react strategically. Since there are capacity constraints for bike production, it is not feasible to put bikes in all targeted cities within a very short timeframe.

Figure 5 plots the β_τ coefficients estimated in the following regression:

$$\begin{aligned}\text{Normalized Bike Placement}_{c,t} &= \alpha_0 \\ &+ \sum_{\tau=-5}^{4} \beta_\tau \cdot \mathbb{1}(t=\tau) \cdot \mathbb{1}(\tau \neq -1) + \beta_5 \cdot \mathbb{1}(t \geq 5) + \delta_c + \mu_t + \varepsilon_{c,t} \quad (2)\end{aligned}$$

In the regression, Normalized Bike Placement$_{c,t}$ is a measure with a range of $[0,1]$, which is defined as $\frac{\text{Bike Placement}_{c,t}}{\text{Maximum Bike Placement in Sample}_c}$, where t corresponds to the number of months after each city's month with the largest bike placement shock. δ_c is city fixed effects, μ_t is year-month fixed effects, and $\varepsilon_{c,t}$ is the error term that varies across cities and over time. The sample period is from May 2017 to January 2020, which avoids later COVID lockdown periods. For each city, the sample only covers periods in which t is not earlier than -5. The figure shows that the magnitude of the largest monthly bike placement shock is large—on average, around 25% of the maximum bike placement of the city during the sample period. Normalized bike placement on average rises by about 10% of the maximum bike placement in the 2 months immediately before the event of the largest monthly bike-placement shock. This pattern of bike placement is consistent with bike-sharing companies' strategic concerns. To address the fierce competition in the bike-sharing industry, once a company decides to enter a city, it is likely to place a lot of bikes in a short period to build up local market power.

At the same time, the timing of the bike placement shock is hard for citizens to predict. Figure 6 shows the monthly time series of the number of cities that are in their month of the largest bike-placement shocks. The critical month for each city's bike placement is distributed broadly over the sample period. This is consistent with the fact that there are capacity constraints on bike production and bike allocation. In that sense, placing shared bikes is like playing chess, in which the players target different cities during different periods. Once they decide on the cities to target, they place many bikes within a very short time frame. Since bike placement is quite staggered and the time of the largest bike placement shock spreads over time, it is hard for citizens of a specific city to predict shocks from the placement of Alipay-bundled shared bikes using only public information.

3 IV Analysis

This section first presents the results of the main specification, which investigates the causal effect of in-person cashless payment flow on BigTech credit provision and consumer take-up of the credit with the IV strategy. It then demonstrates the importance of the payment information channel in facilitating BigTech credit provision. Finally, it illustrates the implications of in-person cashless payment flow for financial inclusion, whereby the causal effects of in-person payment on credit provision mainly hold for the traditionally financially underserved.

3.1 In-person Cashless Payment Flow and Credit Provision

3.1.1 Causal Effects of In-person Cashless Payment on Credit Provision

To analyze how in-person cashless payment flow affects credit provision by BigTech, I estimate the effect using two-stage least squares regressions. In the first stage, the transformed in-person payment flow is instrumented with log-transformed city-level bike placement:

$$g(ipf)_{i,t} = \alpha_1 + \beta_1 \cdot log(bp)_{c,t} + \delta_{1i} + \theta_{1t} + \varepsilon_{1i,t} \tag{3}$$

In the second stage, with the instrumented log-transformed in-person payment flow, I estimate its causal effect on the credit provision variable using the following specification:

$$Y_{i,t} = \alpha_2 + \beta_2 \cdot g(\hat{ipf})_{i,t} + \delta_{2i} + \theta_{2t} + \varepsilon_{2i,t} \tag{4}$$

And the corresponding OLS regression is performed using the following specification:

$$Y_{i,t} = \alpha_0 + \beta_0 \cdot g(ipf)_{i,t} + \delta_{0i} + \theta_{0t} + \varepsilon_{0i,t} \tag{5}$$

where $log(bp)_{c,t}$ is log-transformed bike placement in city c at time t; $g(ipf)_{i,t}$ is the transformed measure of the in-person payment flow of individual i at time t; $g(\hat{ipf})_{i,t}$

is the corresponding instrumented variable; $Y_{i,t}$ is the credit provision variable of individual i at time t; δ_{Ni} $(N = 1, 2, 3)$ represents individual fixed effects; and θ_{Nt} $(N = 1, 2, 3)$ represents year-month fixed effects.

Table 4 shows the results of the regressions specified in equations (3), (4), and (5), in which Panel A reports the estimated effects in the second stage of the two-stage least squares (2SLS) regression, Panel B reports the first-stage results of the 2SLS regression, and Panel C reports the OLS estimates. Columns (1), (2), and (3) focus on the extensive margin, where Credit $\text{Access}_{i,t}$ is a dummy variable that equals 1 if Alipay user i has access to Alipay's virtual credit card at time t and 0 otherwise. Columns (4), (5), and (6) focus on the intensive margin and use only the sample in which users have credit access in the corresponding months, and $log(\text{Credit Access})_{i,t}$ is the log-transformed credit line of Alipay user i's virtual credit card at time t. In columns (1) and (4), $g(ipf)_{i,t}$ is the $log(1 + x)$ transformed in-person payment flow, measured in CNY; in columns (2) and (5), it is the dummy variable that indicates whether the in-person payment flow is positive; in columns (3) and (6), it is the $log(x)$ transformed in-person payment flow, which is only available when the in-person payment flow is positive. All specifications include individual and year-month fixed effects. The granular fixed effects tightly control for heterogeneity across individuals, since the effect of bike placement is identified within each Alipay user. Panel A shows that having positive in-person payment flow in a month leads to a 56.3% increase in the likelihood of gaining credit access for an average Alipay user and a 203.3% increase in the credit line for a user who currently has credit access in the same month. For those who have positive in-person payment flow in the month, a 1% increase in the in-person payment flow leads to a 0.087% increase in the likelihood of gaining credit access for an average Alipay user and a 0.409% increase in the credit line for a user who currently has credit access. Panel B reports both the t-statistic of the estimate in the first stage and the F-statistic of the regression, which indicates that the log-transformed bike placement is a strong instrument. Panel B also reports the estimated correlation between the error terms in the first-stage and second-stage regressions $\rho(u, v)$ (Stock and Yogo, 2005) and the heteroskedastic version of the estimated correlation $\rho(Zu, Zv)$ (Lee et al., 2021). Both show that the correlation is relatively small and the degree of "endogeneity" is low.

Panel C presents the OLS estimates, which are much smaller than the corresponding IV estimates. There are two potential reasons: (1) omitted variables and (2) the nonmonotone payment-credit relationship. First, the OLS estimate can have a downward bias due to omitted variables, when people with less credit based on attributes unobservable to econometricians are more likely to make more in-person cashless payments. The econometric analysis of this issue is illustrated in Section A.1 of the

Appendix. One example of such omitted variables is a negative health shock, which would negatively impact the user's creditworthiness due to a decrease in disposable income and positively affect the in-person payment flow because of spending on treatment and medicine. Second, the nonmonotone relationship between credit provision and in-person cashless payment flow can also lead to a downward bias. Below a certain threshold, more payment flow leads to more information acquisition by the BigTech firm, which in turn facilitates credit provision. However, above the threshold, more payment flow can be regarded as overspending, which causes the borrower to seem riskier and leads to a reduction in the BigTech credit provision. Empirical evidence that supports the nonmonotone relationship is provided in Figure A.4 and Table A.6 of the Appendix.

Table A.7 supports the view that the IV estimates are quite robust, while the OLS estimates tend to be biased. This table presents the results in which city times year-month fixed effects are added, and the interaction between the bike-user indicator and log-transformed bike placement serves as the instrument. This is the Bartik instrument that takes advantage of the different treatments received by bike users and non-bike users. The instrument is very strong and F-statistics are above 300 in all specifications. This setting allows us to add the city times year-month fixed effects, which remove the unobserved time-varying heterogeneity across cities. The IV estimates are quite close to the results in Table 4, while the OLS estimates are much larger than the corresponding results with no city times year-month fixed effects. With better controls and smaller biases, the OLS estimates get closer to the IV estimates.

What is more, the patterns illustrated in Table 4 are very robust under various other settings. Table A.8 shows that the in-person payment flow also affects future credit provision. Table A.9 reports results of regressions that control for the in-person payment flow in the past 1, 2, or 3 months. Table A.10 and Table A.11 further add bike usage and online payment as controls. The effects of the concurrent in-person payment flow on credit provision are still robust, with similar magnitude across different specifications.

3.1.2 Consumer Take-up of BigTech Credit

The credit access and credit line discussed in the previous section are fully determined by the supply side, since no active application is required for Alipay users to use the virtual credit card, and they directly learn about their credit access and credit line by simply checking the account. The real effects of the changes in credit provision also depend on the demand side—that is, consumer take-up of BigTech credit. It is natural to anticipate that more in-person payment flow leads to a higher fraction of spending paid for with the virtual credit card, both in-person and online, for two reasons. The first reason is the learning-by-doing channel, whereby people are more likely to use the virtual credit card when they have more knowledge about Alipay and more trust in Alipay. The second reason is the supply-side channel, whereby Alipay users might use the virtual credit card more frequently when they have a higher credit line.

Results in columns (1) and (2) of Table 5 support the above view of consumer credit take-up. With an exogenous increase in the in-person payment flow, the share paid with Alipay's virtual credit card increases for both in-person payment and online payment. The magnitude of the increase is larger for in-person payment.

Without further assumptions about consumer behavior, the welfare implications of the increased credit take-up are ambiguous. On the one hand, if a consumer is rational, having a more relaxed credit constraint can make her weakly better off with smoother consumption, and the grace period feature of Huabei credit line further increases the credit line's option value. On the other hand, if a consumer has behavioral biases, such as self-control or forecasting problems, she might end up suffering a significant debt burden in the future and reduced welfare. There is supporting evidence for both views and there is a lack of consensus in the consumer finance literature (Zinman, 2015).

The detailed transaction-level data allow me to categorize consumer spending and test whether consumers are more likely to use the digital payment for compulsive spending. If they have self-control problems, more accessible payment methods might make it easier for people to develop addictions. I find no evidence to support this view. Columns (3) and (4) show that the in-person payment flow does not result in a higher fraction of compulsive spending in in-person and online environments.

Both self-control and forecasting problems might cause a consumer to overspend after she get access to a new credit line. I do an event study to evaluate whether consumers have a temporary increase and a sharp reversal in Alipay consumption after they are granted access to a Huabei credit line. Figure A.5 illustrates the changes in Alipay payment flow around the consumer's first month of having access to the Huabei credit line by plotting the β_τ coefficients estimated in the regression:

$$\begin{aligned} log(1 + \text{Total Payment Flow})_{i,t} &= \alpha_0 \\ &+ \sum_{\tau=-5}^{4} \beta_\tau \cdot \mathbb{1}(t = \tau) \cdot \mathbb{1}(\tau \neq -1) + \beta_5 \cdot \mathbb{1}(t \geq 5) + \delta_i + \mu_t + \varepsilon_{i,t} \end{aligned} \quad (6)$$

where $log(1 + \text{Total Payment Flow})_{i,t}$ is the $log(1 + x)$ transformed amount of total payments by individual i at time t using Alipay, t corresponds to the number of months after each individual's first month of having access to a Huabei credit line, δ_i is the individual fixed effects, and μ_t is the year-month fixed effects. For each user with credit access, the sample only covers periods when the t is not earlier than -5. Consumers' Alipay payment reactions to credit access can result from three forces. One is that the consumer has a more relaxed credit constraint, and the credit line allows her to spend more to smooth consumption. The other is that the consumer shifts from other payment instruments, such as cash or debit cards, to Alipay after getting more credit access in Alipay. The third is that the consumer has behavioral biases; thus, she will overspend in the short term and underspend in the long run to pay off the credit balance. We see a temporary increase and a stable long-term shift in the Alipay payment after gaining credit access, but we do not observe a significant reversal in the long run. These findings suggest that the overspending issue might be mild in this setting.

3.2 The Payment Information Channel

3.2.1 Channels for Credit Provision

Two main channels facilitate the credit provided by financial intermediation: the information channel and the enforcement channel. Both information sharing and lender's enforcement power help mitigate information asymmetry problems in the consumer lending market, including adverse selection and moral hazard. These channels can be further classified as follows.

Information channels include:

- Channel 1.1: Use the information on payment flow.
- Channel 1.2: Use the information on credit usage and repayment.
- Channel 1.3: Use the information on the application form.

Enforcement channels include:

- Channel 2.1: Use assets under management (AUM) in the platform as collateral.
- Channel 2.2: Explicitly pledge assets as security for loan repayment.

For banks that do not have borrower payment flow information, the payment flow information channel (channel 1.1) is usually not an option. Instead, the information actively provided by borrowers on the credit application form (channel 1.3) plays an important role before the borrower gains credit access, and the information on credit usage and repayment behaviors (channel 1.2) becomes the most important channel for reducing information asymmetry after the borrower gains credit access. For secured loans such as mortgages, banks usually require borrowers to explicitly pledge the corresponding assets as security for the repayment of loans and forfeit the collateral in the event of default (channel 2.2).

The BigTech company that provides the cashless payment service to borrowers has an advantage in the information flow channel (channel 1.1), whereby the rich information contained in the payment flows reveals valuable information on the borrower's creditworthiness. In the specific setting of Alipay, no application process is required to access a virtual credit card and the explicit pledge of collateral is not an option; thus channels 1.3 and 2.2 are unlikely to play a role in Alipay's credit provision. Instead, information on credit usage and repayment (channel 1.2) can be important, and the enforcement power by the lender can be strong, especially when users use the digital wallets frequently. For example, a borrower's AUM on Alipay's wealth management (channel 2.1) might serve as collateral to facilitate credit provision, since the borrower might worry that the account could be frozen if they do not repay the credit in time.

In this research, I focus on showing the importance of the payment flow information channel (channel 1.1) for credit provision by Alipay and show that the channel holds strongly, whereas channels 1.3 and 2.2 are unavailable and channels 1.2 and 2.1 are controlled for.

3.2.2 Control for the Credit Use and Repayment Information Channel

With the Alipay app, users have several options to make in-person and online payments. On the Alipay platform, they can use the e-wallet account balance, a liquid money market fund called "Yu'ebao," or Huabei, Alipay's virtual credit card. Although Alipay also supports payments using debit card or credit card accounts for some merchants, most transactions on the Alipay platform are paid for with these within-Alipay payment methods since they are cheap, convenient, and widely accepted. I define "in-person credit payment flow" as the amount of in-person Alipay spending paid for by using Alipay's virtual credit card, and this payment flow is directly associated with credit usage and is highly relevant for the credit repayment flow. All other in-person payment flow is defined as the "in-person noncredit payment flow," which does not have a direct relationship with credit usage and repayment.

Table 6 shows the results of the 2SLS and OLS regressions with specifications similar to those of equations (3), (4), and (5), while replacing in-person payment flow with in-person noncredit payment flow, which excludes in-person Alipay payment flow paid for with the virtual credit card. This exclusion helps eliminate the effects of credit use and repayment on BigTech credit provision. Columns (1) and (3) show that the in-person noncredit payment flow has direct effects on BigTech credit provision, and indicates that even after controlling for the credit usage and repayment information channel (channel 1.2), the payment flow information channel (channel 1.1) still matters. However, there might be concerns that the in-person noncredit payment flow is correlated with the in-person credit payment flow, and the specifications in columns (1) and (3) fail to fully exclude the effects of credit usage and repayment. To alleviate concern about the correlation between payment flows, in the specifications of columns (2) and (4) the in-person credit payment flow is added as a control variable in all regressions. The results are still robust, with very close estimates. Moreover, in the second stage of the 2SLS regressions, the in-person credit payment flow does not seem to have a significant impact on credit provision, on either the extensive margin or the intensive margin. The estimated coefficients of the in-person noncredit payment flow measure are larger than those of the in-person payment flow measure in the analysis in Table 4, which indicates that in-person noncredit payment has larger effects than credit payment. This result is reasonable, since the usage of credit directly leads to a heavier repayment burden and riskier consumer profile, while usage of the account balance does not have a direct implication for the risk faced by the BigTech lender.

3.2.3 Control for the Enforcement Channel

Although the explicitly pledged collateral for loan repayment (channel 2.2) is unavailable on the Alipay platform, the user's assets under management (AUM) in Alipay's wealth management products can partially play the role of collateral, since the Alipay platform has the right to freeze the user's account if she does not repay the loan on time. There is a concern that the BigTech credit provision to a user is largely driven by the size of her AUM instead of the information channels. To deal with this concern, the specifications that control for each user's time-varying AUM are analyzed.

Table 7 shows that the relationship between in-person payment flow and BigTech credit provision is robust to adding the AUM variables as controls. Columns (1) and (3) define AUM as all assets in Alipay except for the account balance, while columns (2) and (4) define AUM as all Alipay assets including the account balance. Regardless of which specification is used, the AUM does not have a strong relationship with the credit provision variables, while the in-person payment flow has strong effects on credit provision on both the extensive margin and intensive margin.

3.3 The Financial Inclusion Implications of Cashless Payment

3.3.1 Heterogeneous Outcomes in an Illustrative Example

Before analyzing the heterogeneous effects of cashless payment adoption on credit access empirically, I use an illustrative example to show the potential heterogeneous outcomes predicted by the theory. The detailed setup of the theoretical example is described in Section A.2. Here, I consider the cashless payment adoption as an information shock to the lender and show that it can potentially lead to opposite credit access outcomes for borrowers who are less creditworthy.

There are one lender and a continuum of borrowers in the example. The cashless payment firm, as the only lender, offers a personalized credit line to each borrower. Based on information about the creditworthiness of each borrower, the lender chooses the optimal credit limit to maximize its expected profit. We consider three cases, which represent borrowers' three stages of digital payment adoption. In the first stage—the cash user stage—the borrowers only use cash for transactions, and the lender does not have any creditworthiness information for each borrower. In the second stage—the new digital money adopter stage—borrowers just start to use digital money and submit some personal characteristics information to the lender, and we assume the lender only knows whether the creditworthiness of a borrower is above a threshold or not. In the third stage—the digital money user stage—borrowers start using digital money for daily purchases, which can be observed by the lender, and we assume the lender knows the exact creditworthiness of each borrower.

The relationship between the optimal credit line and the type of borrower in different scenarios is illustrated in Figure 7. Figure (a) shows the financial divide scenario. In this scenario, the threshold value in the new digital money adopter stage is set at 0.25, and some less creditworthy borrowers are worse off in the transition to the digital money user stage. Figure (b) shows the financial inclusion scenario, in which the threshold value in the new digital money adopter stage is instead set at 0.8. As a result, some less creditworthy borrowers are better off in the transition to the digital money user stage. Comparison of the two scenarios shows that better information acquisition by the monopolistic lender does not always lead to more credit access for borrowers with lower creditworthiness.

Without looking into the data, the theory alone does not tell us the impacts of cashless payment adoption on different groups. In the following empirical analysis, we first define traditionally financially underserved segments and then evaluate the heterogeneous effects across segments.

3.3.2 The Traditionally Financially Underserved Segment

My data support the traditional view in China that less educated and older people tend to be financially underserved. Since I do not observe all financial activities of the sampled Alipay users across multiple financial institutions, I use their level of using Alipay financial services as a proxy for their overall financial access. By analyzing their financial behaviors on the Alipay platform, I find that these groups indeed use financial services for fewer activities.

Columns (1), (2), and (3) in Table 8 show the results of cross-sectional regressions that examine the relationship between users' financial activities with Alipay and their personal characteristics. The less educated and the older groups tend to have fewer Alipay financial activities—fewer Alipay-linked debit cards, smaller all-time-high Alipay AUM, and shorter Alipay investment experience. This is consistent with the argument that these groups are less financially literate and are less served by financial institutions.

Less educated and older groups also tend to have lower financial literacy (Lyons et al., 2019), which can potentially further worsen the problem of inadequate access to financial services. My data confirm that this is also a problem for Alipay users who are less educated and older.

Columns (4), (5), and (6) show evidence on how sampled users' education and age relate to measures of financial literacy. Less educated and older users tend to have a smaller likelihood of paying while using their real name, using their own accounts instead of others' accounts, and completing their profile information. These behavioral characteristics are detected automatically by machine learning algorithms. Although it is unclear whether these labels are directly used in making consumer lending decisions for borrowers in the Alipay system, they tend to deliver negative signals about the borrower's creditworthiness, since these behaviors are misaligned with the normal standard.

3.3.3 In-person Cashless Payment and Financial Inclusion

Assuming that different types of data can substitute for each other to improve the ability of financial intermediators to evaluate consumers' credit, the rollout of in-person cashless payment can have financial implications for credit provision. The less educated and the older have previously had less alternative data with which to prove their creditworthiness, and thus they have tended to be underserved by financial intermediation. With an exogenous increase in in-person payment flow by shifting from other payment instruments to Alipay, the marginal increase in the precision of the signal regarding creditworthiness is larger for the previously financially underserved, and it is reasonable to expect that they will benefit more from the shock and gain greater credit access.

Table 9 presents empirical evidence showing the causal relationship between a user's in-person payment flow and the BigTech credit provided to the user, separately for the less and more educated groups and on both the extensive margin and the intensive margin. Panel B shows that, regardless of education group, the first stage is always quite strong. This means that the bike-placement shock consistently increases the in-person cashless payment flow of both the less and more educated. Second-stage results in Panel A reveal that the effects of in-person payment flow on credit provision are quite different for Alipay users with different education levels. The positive relationship only exists for the less educated group and becomes insignificant for the more educated group, on both the extensive margin and the intensive margin. For the less educated group, an increase of in-person payment flow of 1% leads to an increase in the probability of gaining credit access of 0.095% and an increase in the credit line of 0.335%, conditional on credit access. The corresponding numbers for the more educated groups are 0.027% on the extensive margin and 0.035% on the intensive margin, and both estimates are insignificant.

Similarly, the sample can be grouped by age and analyzed separately. Table A.12 shows the corresponding results. Strong first-stage effects hold for both the older and the younger groups. However, in the first stage, there are some differences in the magnitude of effects between age groups. On the extensive margin, the effect of a 1% increase in in-person payment flow on credit access probability is 0.130% for the older group and 0.047% for the younger group, and the former effect is 1.8 times larger. The case is similar on the intensive margin, where the effect of the older group is 1.6 times larger than that of the younger. This is consistent with the previous analysis. The older group has previously been underserved by financial intermediation, and the adoption shock of in-person cashless payment helps them more; thus they end up with larger improvements in credit access.

4 Model-based Analysis

With the IV analysis, we know that an exogenous payment adoption shock leads to more credit provision, and the positive credit provision effects are stronger for the underserved. However, we do not know payment adoption's real effects—e.g. effects on consumer surplus, lender profit, and default rate—which we cannot observe directly in the data. What is more, we can neither say much about why payment data play such an important role in the credit provision nor quantify the information value of payment data. To achieve these goals, I construct and estimate a simple model of a BigTech lender and multiple borrowers with different creditworthiness.

4.1 Model Setup

The setup of the model captures the main features of the BigTech consumer lending market and the product features of the Huabei credit line. In the economy, the cashless payment company acts as the only lender and offers personalized credit lines to borrowers with different creditworthiness. Since we find a strong information channel of payment flow in the empirical analysis, which is almost unchanged when we control for credit repayment or enforcement channels, we assume that the credit evaluation does not rely on credit history or collateral information. Most users of the Huabei credit line are offered a daily interest rate of 0.05%. We assume there is no price discrimination in the economy, and every borrower gets the same interest rate—although the credit limit can vary across different borrowers. We also assume that the lender has sufficient funds, and thus lending more to one borrower does not necessarily mean lending less to others.

Borrowers' consumption gradually shifts from cash to digital money, and we consider three main stages of cashless payment adoption. The first stage is the cash user stage, in which the lender does not know any information about borrowers. The second stage is the new digital money adopter stage, in which the lender knows only the personal characteristics of the borrower, such as age, gender, education, and city of residence. The third stage is the digital money user stage, in which the lender knows the borrower's consumption with the payment data, in addition to their personal characteristics. When we match the data to the model, we assume that at the end of the sample period, all borrowers are in the digital money user stage due to China's fast transition to a cashless society.

There are two periods in the model. In the first period, the lender decides how much to lend to the borrower, and the borrower receives a random income flow and makes borrowing and consumption decisions. In the second period, the borrower receives another random income flow and decides whether to pay off the credit balance or default.

The income flow of borrow i in period $t = 1, 2$ is determined by

$$e_{it} = X_i\beta + y_i + \epsilon_{it} \tag{7}$$

where X_i is a vector of observable characteristics of borrower i, y_i is an unobservable type of borrower i, and ϵ_{it} is an unobservable shock to borrower i in period t. I assume $y_i \in \mathcal{N}(0, \sigma_y^2)$, and thus the density function is $g(y) = \frac{1}{\sigma_y\sqrt{2\pi}}e^{-y^2/2\sigma_y^2}$. I assume idiosyncratic shock $\epsilon_{it} \in \mathcal{N}(0, \sigma_\epsilon^2)$ and $\epsilon_{it} \perp\!\!\!\perp y_i$, and thus the density function is $f(\epsilon) = \frac{1}{\sigma_\epsilon\sqrt{2\pi}}e^{-\epsilon^2/2\sigma_\epsilon^2}$.

The trade-off faced by the lender is the profit from lending and the cost associated with potential default. In period $t = 1$, the lender decides to offer a credit line of l_i to borrower i, and charges a unit fee of R for used credit b_i. In the digital payment era, we assume all consumption is paid with digital money, and the lender observes borrower i's consumption c_i. In period $t = 2$, the lender suffers a loss of the credit line amount l_i if the borrower i defaults.

The lender chooses optimal credit line l_i to maximize its profit:

$$\max_{l_i} R \cdot b_i - \mathsf{E}[\mathbb{1}_i^D | X_i, b_i, c_i; \beta, R, A] \cdot l_i \tag{8}$$

where $\mathbb{1}_i^D$ is a dummy indicating whether borrower i defaults in period $t = 2$.

Borrower i is risk-neutral but heavily discounts the cash flows in period $t = 2$ with a discount rate ρ, and she trades off the utility from consuming in the first period and the disutility from the potential default or credit repayment in the second period. In period $t = 1$, borrower i receives the random income flow e_{i1}, knows about the credit line available to her l_i, decides the amount of credit she would like to use b_i, and makes the consumption c_i. We assume the borrower is hand to mouth in period $t = 1$, and the consumption is $c_i = e_{i1} + (1 - R) \cdot b_i$. In period $t = 2$, borrower i receives the random income flow e_{i2}, and tries to pay off the credit balance b_i with the income and an external illiquid asset A. If the balance cannot be paid off, borrower i defaults and suffers a default cost D.

Borrower i chooses optimal used credit b_i to maximize the utility:

$$\max_{b_i} c_i - \rho \cdot \mathsf{E}[\mathbb{1}_i^D | X_i, b_i, c_i; \beta, R, A] \cdot D - \rho \cdot (1 - \mathsf{E}[\mathbb{1}_i^D | X_i, b_i, c_i; \beta, R, A]) \cdot b_i \tag{9}$$

such that

$$0 \leq b_i \leq l_i$$

Borrower i always knows her own consumption, regardless of whether she uses cash or digital money for daily purchases. The lender would not know borrower i's consumption if she is a cash user, and would know if she is a digital money user. When the consumption is in the information set, the expected default probability would be

$$\begin{aligned}
&\mathsf{E}[\mathbb{1}_i^D|X_i, b_i, c_i; \beta, R, A] \\
&= \mathsf{E}[\mathbb{1}(X_i\beta + y_i + \epsilon_{i1} - \epsilon_{i1} + \epsilon_{i2} + A - b_i < 0)|X_i, b_i, c_i; \beta, R, A] \\
&= \int_{-\infty}^{+\infty} \mathbb{1}(\epsilon_{i2} - \epsilon_{i1} < b_i - A - e_{i1}|b_i, e_{i1}) f(\epsilon_{i1}) f(\epsilon_{i2}) d\epsilon_{i1} d\epsilon_{i2} \\
&= \Phi(\frac{b_i - A - e_{i1}}{\sqrt{2}\sigma_\epsilon})
\end{aligned} \tag{10}$$

which does not depend on X_i, the observed personal characteristics.

Solving the lender's problem, we get the following first-order condition, which helps identify the external illiquid asset A:

$$R \cdot \frac{\partial b_i}{\partial l_i} - \Phi(\frac{b_i - A - e_{i1}}{\sqrt{2}\sigma_\epsilon}) - \phi(\frac{b_i - A - e_{i1}}{\sqrt{2}\sigma_\epsilon}) \cdot \frac{l_i}{\sqrt{2}\sigma_\epsilon} \cdot \frac{\partial b_i}{\partial l_i} = 0 \tag{11}$$

Solving borrower i's problem, we get the following first-order condition, which helps identify the borrower's default cost D:

$$(1 - R) - \phi(\frac{b_i - A - e_{i1}}{\sqrt{2}\sigma_\epsilon}) \cdot \frac{\rho \cdot (D - b_i)}{\sqrt{2}\sigma_\epsilon} - \rho \cdot [1 - \Phi(\frac{b_i - A - e_{i1}}{\sqrt{2}\sigma_\epsilon})] = 0 \tag{12}$$

4.2 Model Estimation

I clean the data by winsorizing consumption and used credit at 5% and 95%, and drop the months with zero consumption. The parameters I will estimate with data are $\{\beta, \sigma_\epsilon, \sigma_y, A, D\}$, and the parameters I plan to calibrate are $\{R, \rho\}$. The structure of the model allows me to pin down these parameters one by one with the following estimation procedure.

First, I calibrate credit usage fee $R = 0.03$ and the discounting parameter $\rho = 0.9$.

Second, I assume borrower i has fully shifted from cash to digital money for consumption when her credit line stops increasing. In these months, $c_i = e_{i1} + (1 - R) \cdot b_i$ holds, and the monthly income can be backed up with the consumption and used credit. Since monthly income is determined by $e_{i1} = X_i\beta + y_i + \epsilon_{i1}$, the variations in monthly income within individuals help us to estimate σ_ϵ. I use the average monthly consumption, used credit, and income as the observed values c_i, b_i, and e_{i1}, respectively. Table A.13 reports the distribution of these key variables.

Third, I estimate the parameters β and σ_y by running the regression $e_{i1} = X_i\beta + y_i + \epsilon_{i1}$. In the estimation, the observables X_i include gender, education, age, and city, and the specification is

$$e_{i1} = \beta_0 + \beta_{male} \cdot \mathbb{1}_i^{male} + \sum_{edu \in U} \beta_{edu} \cdot \mathbb{1}_i^{edu} + \sum_{k \in K} \beta_k \cdot \mathbb{1}_i^{k-5<age\leq k} + \sum_{city \in C} \beta_{city} \cdot \mathbb{1}_i^{city} + u_i \tag{13}$$

where $U = \{\text{Below College}, \text{Undergraduate}, \text{Graduate}\}$; $K = \{1930, 1935, ..., 2010\}$; C includes 340 unique cities in China; and error term $u_i = y_i + \epsilon_{i1}$—thus $u_i \in \mathcal{N}(0, \sigma_y^2 + \sigma_\epsilon^2)$.

Fourth, I estimate the external illiquid asset A by using lender's FOC as the moment condition. I assume the lender uses heuristics to predict used credit: $b_i = \lambda \cdot l_i$.

Finally, I estimate D by using the borrower's FOC as the moment condition.

With the estimation procedure, I get the following estimated parameter values:

Parameter	Value	Description
σ_ϵ	864.8	Standard deviation of unobservable idiosyncratic income shocks
σ_y	1,344.0	Standard deviation of unobservable type of borrowers
A	4,692.0	External funding that can be used to pay off the credit balance
D	57,039.7	Utility cost to a borrower if she defaults in the second period
R^2_{ols}	0.0807	R squared of the OLS regression that predicts income

The model yields a prediction for the equilibrium credit line offered to each borrower. In the cross-section, the predicted credit lines explain 12% of variation in the data, and the relationship between the observed credit line and the model predicted credit line is

$$\text{Credit Line}_i^{observed} = \underset{(89.81)}{1777.70} + \underset{(0.01)}{0.94} \cdot \text{Credit }\widehat{\text{Line}}_i^{cashless}$$

where the slope is close to unity.

4.3 Counterfactual Analysis

We are interested in the information value of payment data, and thus the key counterfactual is the new digital money user case, in which borrowers still borrow from the lender, but consume with cash. In this case, the lender lends to borrowers without knowing their consumption, and evaluates the creditworthiness of borrowers with only observed personal characteristics. We simulate the new digital money user case as the counterfactual to compare steady states with the simulated digital money user case.

Since the lender can't see borrower i's consumption c_i in the new digital money user case, the expected default probability becomes

$$
\begin{aligned}
&\mathsf{E}[\mathbb{1}_i^D | X_i, b_i; \beta, R, A] \\
&= \mathsf{E}[\mathbb{1}(X_i\beta + y_i + \epsilon_{i2} + A - b_i < 0) | X_i, b_i; \beta, R, A] \\
&= \int_{-\infty}^{+\infty} \mathbb{1}(y_i + \epsilon_{i2} < b_i - A - X_i\beta | X_i, b_i; \beta) g(y) f(\epsilon_{i2}) dy d\epsilon_{i2} \\
&= \Phi(\frac{b_i - A - X_i\beta}{\sqrt{\sigma_\epsilon^2 + \sigma_y^2}})
\end{aligned}
\tag{14}
$$

which depends on X_i, the observed personal characteristics, unlike Equation (10).

Table 10 uses two panels to compare the steady states between the simulated digital money user case and the simulated counterfactual, which is the new digital money adopter case. Panel A focuses on the average effects, and Panel B highlights the heterogeneous effects across education, age, and gender. The only difference between the two cases is whether the lender knows borrowers' consumption, in addition to observed personal characteristics; thus the comparison can help quantify the information value of the payment data.

Panel A shows that the payment data can increase the average credit line from 3,619.9 CNY to 5,707.5 CNY, which is a 57.7% increase. It also increases used credit, consumer welfare, and lender profit. The monthly total welfare increase is 17.8 CNY on average, which is roughly 3 USD. Multiplying this number by 1 billion users, we get a welfare improvement value of roughly 36 billion USD per year. The annualized default rate also increases in the transition to the digital money user case.

Panel B shows that changes in the credit line, consumer welfare, lender profit, and default rate are not homogeneous. They seem to be larger for the less educated and the older, which is consistent with the findings in the IV analysis. These analyses suggest that the payment information leads to better financial inclusion, since the less educated and the older not only get a relatively higher credit line after the digital payment shock but also get relatively higher consumer surplus while generating more profit for the lender.

5 Conclusion

The easy adoption process, high convenience, and low intermediation fee all contribute to the success of the in-person cashless payment in China. Since using cashless payment in the in-person environment is not very different from using cash for daily purchases, the extremely low barrier makes the technology accessible even to those who were previously financially unserved or underserved. In the transition from a cash economy to a cashless economy, users naturally accumulate their payment records while using digital payment services. This paper shows that payment data can serve as valuable digital assets that facilitate credit provision to the relatively disadvantaged.

By using deidentified data from Alipay, the world's leader in mobile payment with 1 billion active users, I document that an exogenous increase in the in-person cashless payment flow leads to more credit provision by financial intermediation. This increase in credit provision results from the useful information for credit evaluation provided by the payment flow. The information goes beyond what is available from credit usage, repayment, and assets under management. I use a novel instrument by taking advantage of the staggered placement of Alipay-bundled dockless shared bikes across cities to alleviate endogeneity concerns and conduct several tests to confirm the instrument's validity.

I find that the previously financially underserved benefit more from mobile payment adoption, and I propose a simple theoretical framework to provide insight into the underlying forces that can generate the corresponding predictions. I also estimate a simple model to quantify the information value of payment data.

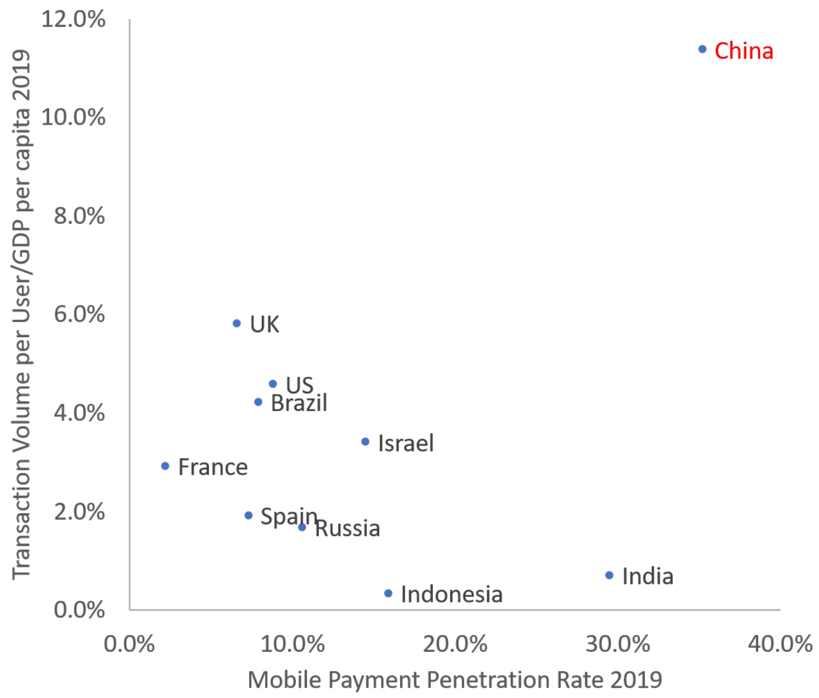
These findings have strong policy implications: The prevalence of mobile phone adoption can potentially provide new opportunities for financial inclusion, and mobile payment can support a sustainable business model for lending to the poor. With the remarkably rapid development of mobile payment in China, it is possible that other developing countries can also experience abrupt changes in the cashless payment market in the future. Once that occurs, the digital payment system can function as an infrastructure for credit evaluation and credit provision.

Note that an increase in credit provision to the relatively underserved does not mean that it is optimal for financial intermediation to lend to everyone who has payment data. For instance, it might not be profitable to lend to the extremely disadvantaged. In these cases, the government could potentially subsidize individuals with a fiscal transfer. This paper makes a start in studying the implications of digital payments in the consumer credit market. Much more study is necessary to understand the welfare implications of public policies.

Figure 1: Mobile Payment Penetration across Countries

These figures show the GDP-adjusted mobile payment transaction volume per user and the mobile payment penetration rate for selected countries in 2019 and 2021. Data sources are the Statista Digital Market Outlook and the World Bank.

(a) 2019



(b) 2021

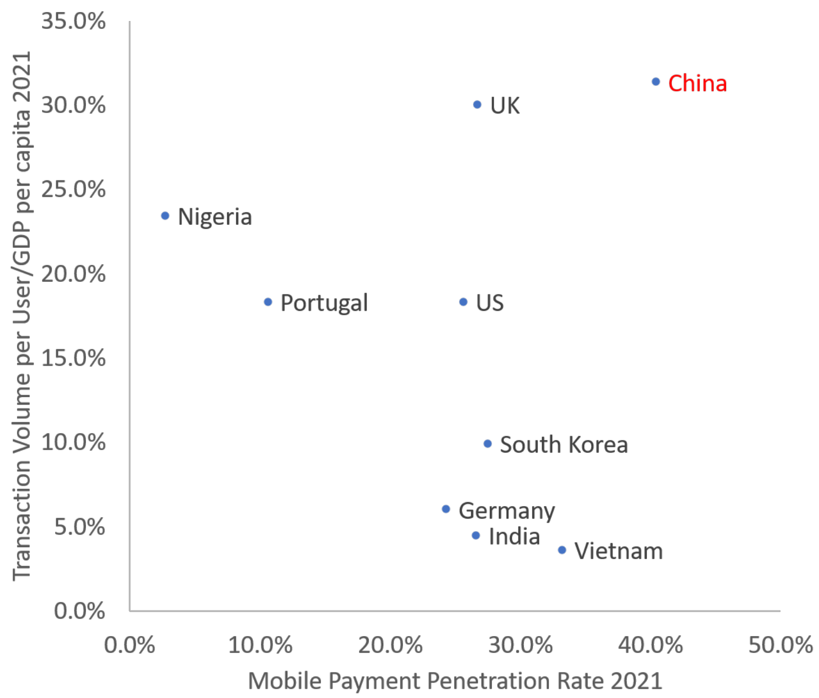


Figure 2: Transaction Volume of Mobile and Card Payment in China and US

This figure presents the time series of the GDP-adjusted transaction volume of mobile and card payments in China and the US from 2012 to 2018. Data sources are the US Federal Reserve, the People's Bank of China (PBOC), and the World Bank.

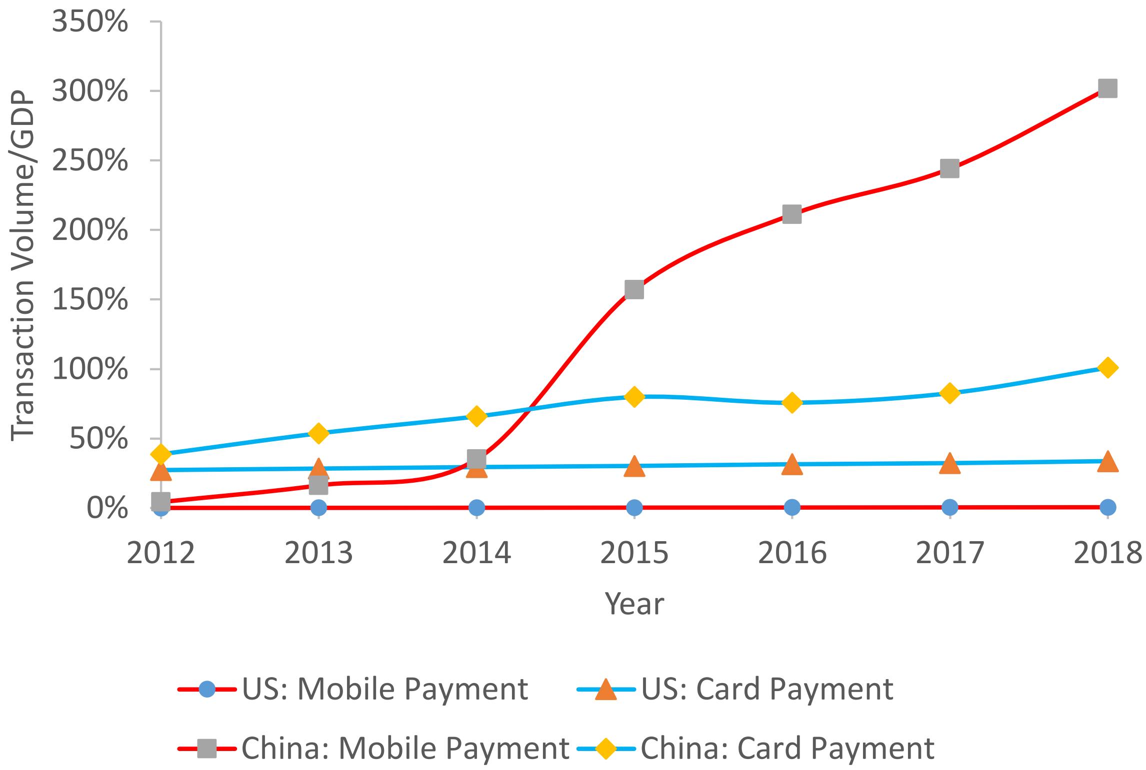


Figure 3: Logic Flow of the Instrumental Variable

This figure presents a graphical illustration of the mechanisms that show how the city-wide placement of Alipay-bundled shared bikes affects the city's residents' in-person Alipay payment at the individual level.

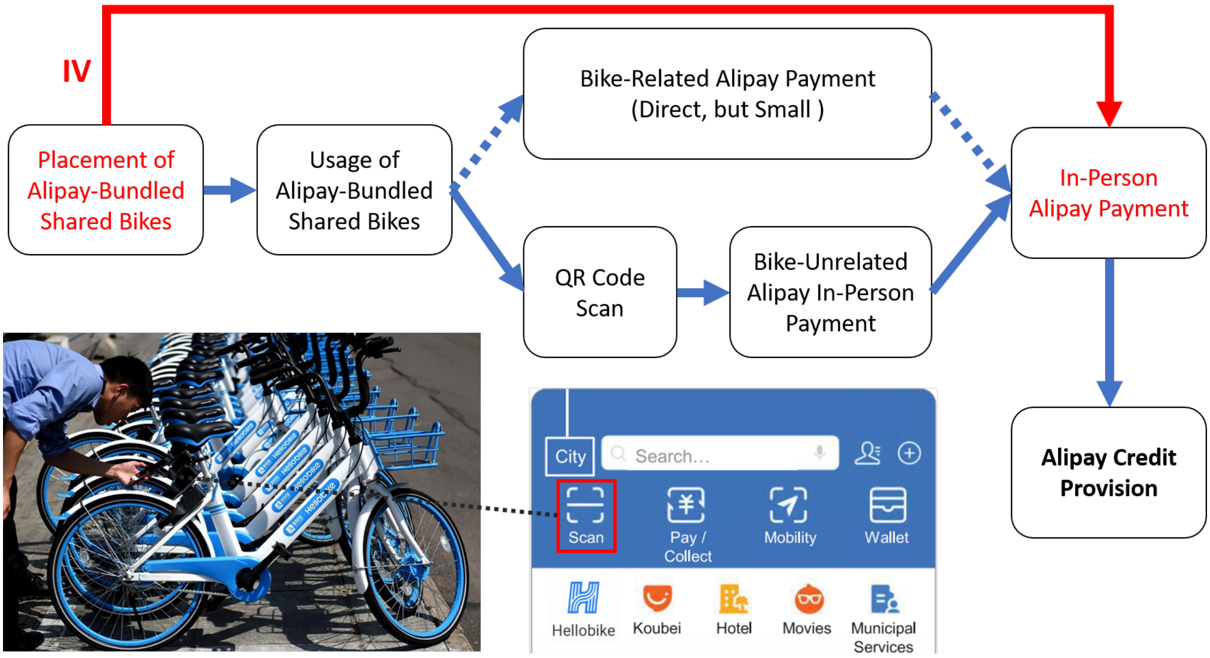


Figure 4: Bike Adoption and Non-bike Payment Flow

This figure plots the β_τ coefficients estimated in the following regression:

$$log(1 + \text{In-Person Non-Bike Payment Flow})_{i,t} = \alpha_0 + \sum_{\tau=-5}^{4} \beta_\tau \cdot \mathbb{1}(t = \tau) \cdot \mathbb{1}(\tau \neq -1) + \beta_5 \cdot \mathbb{1}(t \geq 5) + \delta_i + \mu_t + \varepsilon_{i,t}$$

where $log(1 + \text{In-Person Non-Bike Payment Flow})_{i,t}$ is the $log(1 + x)$ transformed amount of in-person payments on purchases not directly related to the usage of shared bikes by individual i at time t using Alipay, t corresponds to the number of months after each individual's month of the first usage of Alipay-bundled shared bikes, δ_i is the individual fixed effects, μ_t is the year-month fixed effects, and $\varepsilon_{i,t}$ is the error term that varies across individuals and over time. The sample covers only users who used the Alipay-bundled shared bikes at least once in the sample period, which is from May 2017 to September 2020. For each bike user, the sample only covers periods in which the t is not earlier than -5.

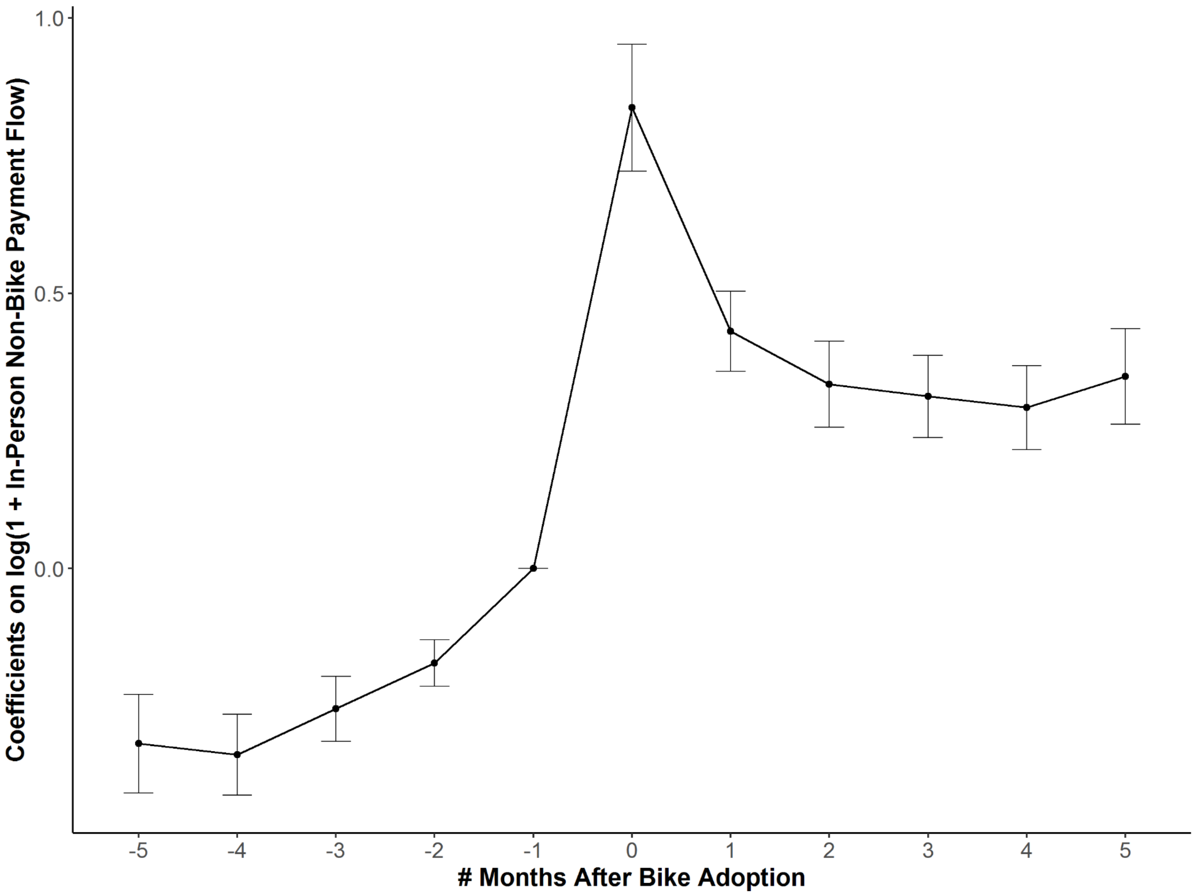


Figure 5: Staggered Placement of Shared Bikes

This figure plots the β_τ coefficients estimated in the following regression:

$$\text{Normalized Bike Placement}_{c,t} = \alpha_0 + \sum_{\tau=-5}^{4} \beta_\tau \cdot \mathbb{1}(t=\tau) \cdot \mathbb{1}(\tau \neq -1) + \beta_5 \cdot \mathbb{1}(t \geq 5) + \delta_c + \mu_t + \varepsilon_{c,t}$$

where Normalized Bike Placement$_{c,t}$ is defined as $\frac{\text{Bike Placement}_{c,t}}{\text{Maximum Bike Placement in Sample}_c}$, which is a measure with a range of $[0,1]$, t corresponds to the number of months after each city's month with the largest bike placement shock, δ_c is the city fixed effects, μ_t is the year-month fixed effects, and $\varepsilon_{c,t}$ is the error term that varies across cities and over time. The sample period is from May 2017 to January 2020, which avoids the later COVID lockdown periods. For each city, the sample only covers periods in which the t is not earlier than -5.

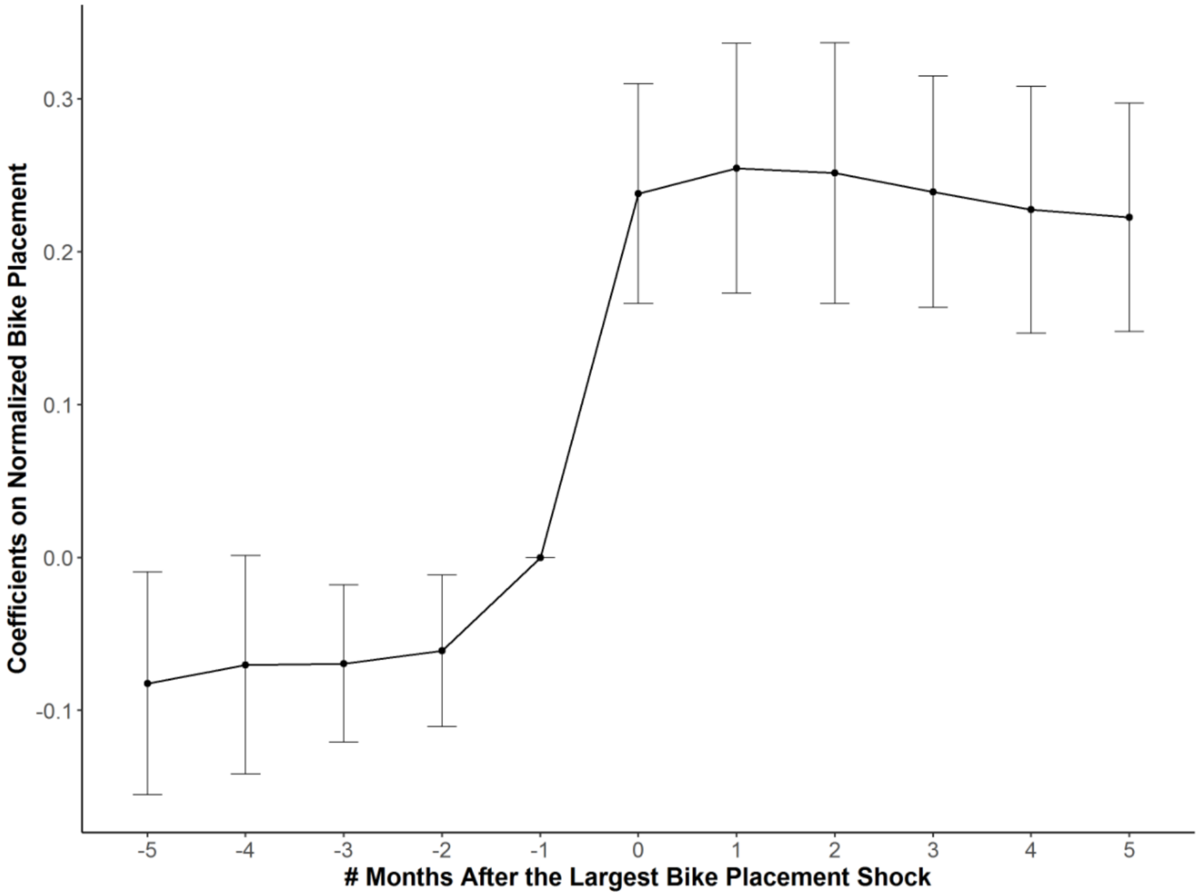


Figure 6: Broad Distribution of Bike-placement Shock

This figure describes the number of cities in the month of their largest bike-placement shock in the period from May 2017 to January 2020, before the later COVID lockdown periods.

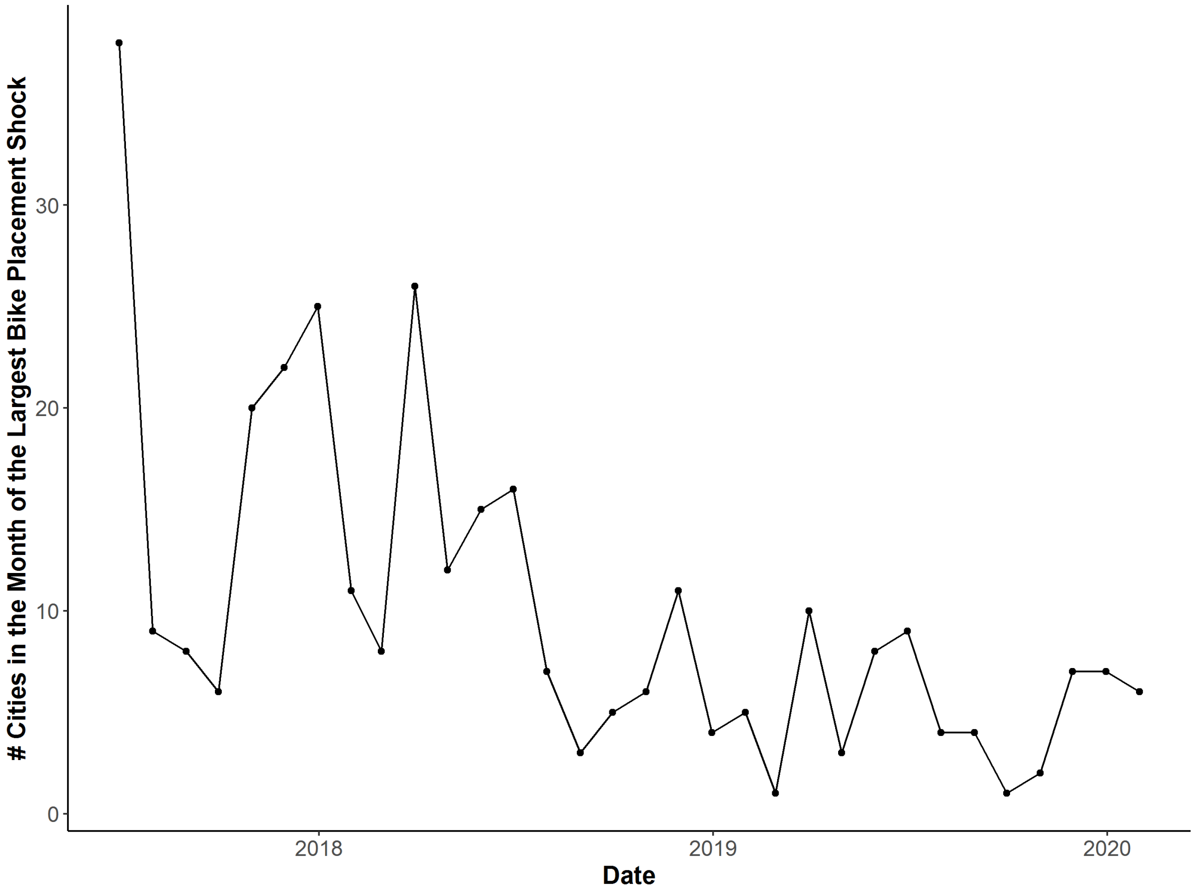
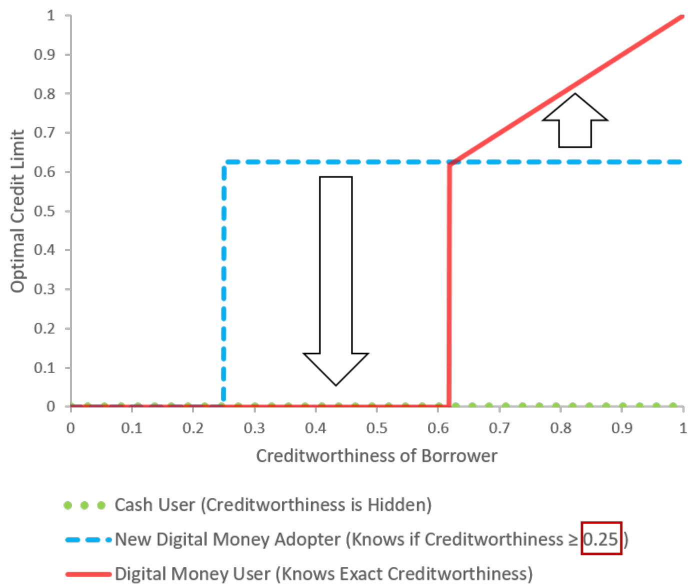


Figure 7: Cashless Payment's Financial Inclusion Implication: An Example

The figures illustrate how a lender sets optimal credit limits for heterogeneous borrowers in an economy with a continuum of borrowers. The illustrative example's setup is in Section A.2. Given knowledge about the borrower's type, the lender chooses the optimal credit limit to maximize the expected profit. There are three cases with different information sets for the lender based on the borrower's digital payment adoption stage. In stage one (green dotted line), borrowers use cash and the lender knows nothing about their creditworthiness. In stage two (blue dash line), borrowers adopt digital money and the lender knows if their creditworthiness is above a threshold. In stage three (red solid line), borrowers use digital money and the lender knows their exact creditworthiness. Figure (a) shows the financial divide scenario, where some less creditworthy borrowers lose in the transition from stage two to three. Figure (b) shows the financial inclusion scenario, where some less creditworthy borrowers gain in the transition to stage three.

(a) Scenario of ***Financial Divide***



(b) Scenario of ***Financial Inclusion***

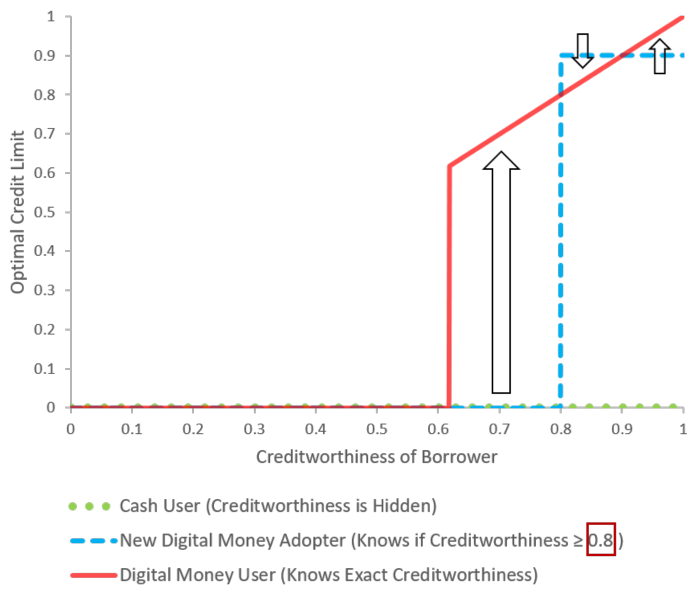


Table 1: Summary Statistics

The table summarizes the key variables for our analysis of 41,485 Alipay users from May 2017 to September 2020. The variables are categorized into three types at different levels. At the individual level, # Active months$_i$ counts the months with payment activities; Is Male$_i$ is a dummy for male; Low Education$_i$ is a dummy for below bachelor's degree; Birth Year$_i$ records the birth year; Bike User$_i$ is a dummy for using a shared bike at least once. At the city-month level, $log(\text{Bike Placement})_{c,t}$ is the log number of active shared bikes in city c at time t. At the individual-month level, Credit Access$_{i,t}$ is a dummy for having access to Alipay's virtual credit card at time t; $log(\text{Credit Line})_{i,t}$ is the log credit line conditional on Credit Access$_{i,t}$; $log(\text{In-Person Payment Flow})_{i,t}$ and $log(\text{Online Payment Flow})_{i,t}$ are the log amounts of in-person and online payments using Alipay; Virtual Credit Card Share in In-Person Payment$_{i,t}$ and Virtual Credit Card Share in Online Payment$_{i,t}$ are the shares of in-person and online payments using the virtual credit card; Compulsive Spending Share in In-Person Payment$_{i,t}$ measures the share of in-person Alipay payments made by individual i at time t for cigarettes, games, and lotteries; Compulsive Spending Share in Online Payment$_{i,t}$ measures the share of online Alipay payments made by individual i at time t for cigarettes, games, lotteries, or live streaming services.

	N	Mean	Std	Min	p25	Median	p75	Max
Individual Level								
# Active months$_i$	41,485	31.86	11.38	1.00	24.00	37.00	41.00	41.00
Is Male$_i$	41,214	0.54	0.50	0.00	0.00	1.00	1.00	1.00
Low Education$_i$	41,459	0.88	0.33	0.00	1.00	1.00	1.00	1.00
Birth Year$_i$	41,214	1,983.38	12.75	1,930.00	1,974.00	1,985.00	1,993.00	2,014.00
Bike User$_i$	41,485	0.29	0.45	0.00	0.00	0.00	1.00	1.00
City-Month Level								
$log(\text{Bike Placement})_{c,t}$	12,665	7.08	3.39	0.00	4.11	7.85	9.91	13.91
Individual-Month Level								
Credit Access$_{i,t}$	1,321,837	0.62	0.49	0.00	0.00	1.00	1.00	1.00
$log(\text{Credit Line})_{i,t}$	819,812	7.88	1.58	3.00	6.91	8.13	9.13	11.02
$log(\text{In-Person Payment Flow})_{i,t}$	688,428	5.70	2.29	-4.61	4.31	6.04	7.27	15.88
$log(\text{Online Payment Flow})_{i,t}$	843,993	5.76	1.80	-4.61	4.70	5.88	6.93	15.74
Virtual Credit Card Share in In-Person Payment$_{i,t}$	688,428	0.34	0.42	0.00	0.00	0.04	0.82	1.00
Virtual Credit Card Share in Online Payment$_{i,t}$	843,993	0.33	0.41	0.00	0.00	0.01	0.80	1.00
Compulsive Spending Share in In-Person Payment$_{i,t}$	688,428	0.03	0.14	0.00	0.00	0.00	0.00	1.00
Compulsive Spending Share in Online Payment$_{i,t}$	843,993	0.01	0.10	0.00	0.00	0.00	0.00	1.00

Table 2: Effects of Bike Placement on Payment and Credit

These tables report the effects of city-level placement of shared bikes on the individual-level in-person payment flow and digital credit access. $log(\text{Bike Placement})_{c,t}$ is the log number of active shared bikes in city c at time t. Bike User$_i$ is a dummy for using a shared bike at least once from May 2017 to September 2020. After First Bike Usage$_{i,t}$ is a dummy for after the first bike usage. $log(1 + \text{In-Person Payment Flow})_{i,t}$ and $log(1 + \text{Credit Line})_{i,t}$ are the $log(1 + x)$ transformed amounts of in-person payment flow and credit line through Alipay in CNY. Panel A shows the effects on payment flow and Panel B shows the effects on credit line. Columns (1) and (2) use individual and year-month fixed effects with the full sample. Column (3) adds city times year-month fixed effects with the bike user sample. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

Panel A. Bike Placement and Individual-level In-person Payment Flow

	$log(1 + \text{In-Person Payment Flow})_{i,t}$		
	(1)	(2)	(3)
$log(\text{Bike Placement})_{c,t}$	0.041***	0.011	
	(0.010)	(0.009)	
Bike User$_i$ $\times$ $log(\text{Bike Placement})_{c,t}$		0.103***	
		(0.017)	
After First Bike Usage$_{i,t}$			-0.123
			(0.161)
After First Bike Usage$_{i,t}$ $\times$ $log(\text{Bike Placement})_{c,t}$			0.049***
			(0.014)
Individual FE	YES	YES	YES
Year-Month FE	YES	YES	-
City $\times$ Year-Month FE	NO	NO	YES
Clustered by City and Year-Month	YES	YES	YES
Sample	Full Sample	Full Sample	Bike Users
Observations	1,238,309	1,238,309	435,872
Adjusted R^2	0.551	0.552	0.490
Note:			*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Panel B. Bike Placement and Individual-level Digital Credit Line

	$log(1 + \text{Credit Line})_{i,t}$		
	(1)	(2)	(3)
$log(\text{Bike Placement})_{c,t}$	0.027***	0.009	
	(0.008)	(0.010)	
Bike User$_i$ $\times$ $log(\text{Bike Placement})_{c,t}$		0.060**	
		(0.023)	
After First Bike Usage$_{i,t}$			-0.231
			(0.157)
After First Bike Usage$_{i,t}$ $\times$ $log(\text{Bike Placement})_{c,t}$			0.070***
			(0.013)
Individual FE	YES	YES	YES
Year-Month FE	YES	YES	-
City $\times$ Year-Month FE	NO	NO	YES
Clustered by City and Year-Month	YES	YES	YES
Sample	Full Sample	Full Sample	Bike Users
Observations	1,238,309	1,238,309	435,872
Adjusted R^2	0.800	0.800	0.774
Note:			*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 3: Bike-usage Intensity and Heterogeneous Bike-placement Effects

This table reports the heterogeneous effects of city-level placement of shared bikes on the individual-level in-person payment flow and digital credit for non-bike users, one-time bike users, and repeat bike users. $log(\text{Bike Placement})_{c,t}$ is the log number of active shared bikes in city c at time t. One-Time Bike User$_i$ and Repeat Bike User$_i$ are dummies for using a shared bike exactly once and at least twice from May 2017 to September 2020. $log(1 + \text{In-Person Payment Flow})_{i,t}$ and $log(1 + \text{Credit Line})_{i,t}$ are the $log(1 + x)$ transformed amounts of in-person payment flow and credit line through Alipay in CNY. Columns (1) and (3) use individual and year-month fixed effects; columns (2) and (4) add city times year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	$log(1 + \text{In-Person Payment Flow})_{i,t}$		$log(1 + \text{Credit Line})_{i,t}$	
	(1)	(2)	(3)	(4)
$log(\text{Bike Placement})_{c,t}$	0.011		0.009	
	(0.009)		(0.010)	
One-Time Bike User$_i$ $\times$ $log(\text{Bike Placement})_{c,t}$	0.088***	0.072***	0.048**	0.035
	(0.020)	(0.019)	(0.023)	(0.025)
Repeat Bike User$_i$ $\times$ $log(\text{Bike Placement})_{c,t}$	0.106***	0.078***	0.062**	0.040
	(0.018)	(0.017)	(0.025)	(0.029)
Individual FE	YES	YES	YES	YES
Year-Month FE	YES	-	YES	-
City $\times$ Year-Month FE	NO	YES	NO	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Observations	1,238,309	1,238,309	1,238,309	1,238,309
Adjusted R^2	0.552	0.555	0.800	0.801

Note: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 4: In-person Payment Flow and Credit Provision

The table shows the robust relationship between in-person payment flow and BigTech credit with different specifications on both margins. Credit $\text{Access}_{i,t}$ is a dummy for having access to Alipay's virtual credit card at time t. $log(\text{Credit Line})_{i,t}$ is the log credit line conditional on Credit $\text{Access}_{i,t}$. Measure of In-Person Payment $\text{Flow}_{i,t}$ is the amount of in-person payment flow through Alipay at time t, defined differently in different columns: $log(1 + \text{In-Person Payment Flow})_{i,t}$ in CNY in columns (1) and (4); a dummy for positive In-Person Payment $\text{Flow}_{i,t}$ in columns (2) and (5); $log(\text{In-Person Payment Flow})_{i,t}$ conditional on positive In-Person Payment $\text{Flow}_{i,t}$ in columns (3) and (6). $log(\text{Bike Placement})_{c,t}$ is the log number of active shared bikes in city c at time t, conditional on positive $log(\text{Bike Placement})_{c,t}$. Panel A reports 2SLS estimates, instrumenting for payment flow with bike placement; Panel B reports the first stage; Panel C reports OLS estimates. All columns use individual and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit $\text{Access}_{i,t}$			$log(\text{Credit Line})_{i,t}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Two-Stage Least Squares						
Measure of In-Person Payment $\text{Flow}_{i,t}$	0.086***	0.563***	0.087**	0.281***	2.033**	0.409***
	(0.024)	(0.175)	(0.043)	(0.085)	(0.766)	(0.132)
Panel B. First Stage for Measure of In-Person Payment $\text{Flow}_{i,t}$						
$log(\text{Bike Placement})_{c,t}$	0.041***	0.006***	0.030***	0.043***	0.006***	0.024***
	(0.010)	(0.002)	(0.009)	(0.012)	(0.002)	(0.008)
F-Statistic	15.5	10.8	11.2	13.9	10.6	9.1
Estimated $\rho(u, v)$	0.088	0.092	0.047	0.075	0.038	0.080
Estimated $\rho(Zu, Zv)$	0.091	0.093	0.049	0.078	0.042	0.079
Adjusted R^2	0.551	0.465	0.432	0.527	0.439	0.401
Panel C. Ordinary Least Squares						
Measure of In-Person Payment $\text{Flow}_{i,t}$	0.010***	0.062***	0.007***	0.022***	0.072***	0.029***
	(0.001)	(0.007)	(0.001)	(0.003)	(0.023)	(0.002)
Adjusted R^2	0.740	0.741	0.700	0.836	0.835	0.841
Form of the IPF Measure	$log(1+x)$	$\mathbb{1}(x>0)$	$log(x)$	$log(1+x)$	$\mathbb{1}(x>0)$	$log(x)$
Individual FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Full Sample	Has Credit	Has Credit	Has Credit
Observations	1,238,309	1,238,309	662,010	779,283	779,283	516,570

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 5: In-person Payment Flow and Consumer Behavior

This table presents empirical evidence showing the causal relationship between a user's in-person payment flow and the structure of the payment flows, in both the in-person payment and the online payment settings. Virtual Credit Card Share$_{i,t}$ is the share of Alipay payments using the virtual credit card by individual i at time t. Compulsive Spending Share$_{i,t}$ is the share of Alipay payments on cigarettes, games, lotteries, or live streaming services by individual i at time t. $log(1 +$ In-Person Payment Flow$)_{i,t}$ is the $log(1 + x)$ transformed amount of in-person payment flow through Alipay in CNY. log(Bike Placement$)_{c,t}$) is the log number of active shared bikes in city c at time t. Panel A reports 2SLS estimates, instrumenting for payment flow with bike placement; Panel B reports the first stage; Panel C reports OLS estimates. All columns use individual and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Virtual Credit Card Share$_{i,t}$		Compulsive Spending Share$_{i,t}$	
	In-Person Payment	Online Payment	In-Person Payment	Online Payment
	(1)	(2)	(3)	(4)
Panel A. Two-Stage Least Squares				
$log(1 +$ In-Person Payment Flow$)_{i,t}$	0.094***	0.030***	0.004	0.002
	(0.034)	(0.011)	(0.010)	(0.002)
Panel B. First Stage for $log(1 +$ In-Person Payment Flow$)_{i,t}$				
log(Bike Placement$)_{c,t}$	0.028***	0.064***	0.028***	0.064***
	(0.009)	(0.014)	(0.009)	(0.014)
F-Statistic	11.0	22.7	11.0	22.7
Adjusted R^2	0.434	0.505	0.434	0.505
Panel C. Ordinary Least Squares				
$log(1 +$ In-Person Payment Flow$)_{i,t}$	-0.009***	0.008***	0.0002	-0.0003***
	(0.002)	(0.001)	(0.000)	(0.000)
Adjusted R^2	0.472	0.480	0.216	0.222
Individual FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Observations	662,010	806,938	662,010	806,938
Note:				*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 6: In-person Noncredit Payment Flow and Credit Provision

The table shows the causal relationship between in-person noncredit payment flow and BigTech credit on both margins. Credit Access$_{i,t}$ is a dummy for having access to Alipay's virtual credit card at time t. $log(\text{Credit Line})_{i,t}$ is the log credit line conditional on Credit Access$_{i,t}$. $log(1+\text{In-Person Noncredit Payment Flow})_{i,t}$ and $log(1+\text{In-Person Credit Payment Flow})_{i,t}$ are the $log(1+x)$ transformed amounts of in-person Alipay payment flow without and with the virtual credit card in CNY. $log(\text{Bike Placement})_{c,t})$ is the log number of active shared bikes in city c at time t. Panel A reports 2SLS estimates, instrumenting for noncredit payment flow with bike placement; Panel B reports the first stage; Panel C reports OLS estimates. All columns use individual and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit Access$_{i,t}$		$log(\text{Credit Line})_{i,t}$	
	(1)	(2)	(3)	(4)
Panel A. Two-Stage Least Squares				
$log(1+\text{In-Person Noncredit Payment Flow})_{i,t}$	0.094***	0.095**	0.329***	0.358***
	(0.024)	(0.026)	(0.103)	(0.124)
$log(1+\text{In-Person Credit Payment Flow})_{i,t}$		-0.005		-0.044
		(0.006)		(0.029)
Panel B. First Stage for $log(1+\text{In-Person Noncredit Payment Flow})_{i,t}$				
$log(\text{Bike Placement})_{c,t}$	0.037***	0.035***	0.037***	0.031***
	(0.009)	(0.009)	(0.010)	(0.009)
$log(1+\text{In-Person Credit Payment Flow})_{i,t}$		0.218***		0.230***
		(0.009)		(0.007)
F-Statistic	16.1	16.1	13.4	12.6
Adjusted R^2	0.475	0.492	0.457	0.480
Panel C. Ordinary Least Squares				
$log(1+\text{In-Person Noncredit Payment Flow})_{i,t}$	0.006***	0.004***	0.003	-0.004*
	(0.001)	(0.001)	(0.002)	(0.002)
$log(1+\text{In-Person Credit Payment Flow})_{i,t}$		0.015***		0.039***
		(0.001)		(0.003)
Adjusted R^2	0.739	0.742	0.835	0.837
Individual FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Has Credit	Has Credit
Observations	1,238,309	1,238,309	779,283	779,283
Note:		*$p<0.1$; **$p<0.05$; ***$p<0.01$		

Table 7: In-person Payment Flow and Credit Provision, with Enforcement Controls

The table shows the causal relationship between in-person payment flow and BigTech credit after controlling for time-varying assets under management (AUM) on both margins. Credit Access$_{i,t}$ is a dummy for having access to Alipay's virtual credit card at time t. $log(\text{Credit Line})_{i,t}$ is the log credit line conditional on Credit Access$_{i,t}$. $log(1 + \text{In-Person Payment Flow})_{i,t}$ and $log(1 + \text{Assets under Management})_{i,t}$ are the $log(1 + x)$ transformed amounts of in-person payment flow and AUM on Alipay's platform in CNY. $log(\text{Bike Placement})_{c,t}$) is the log number of active shared bikes in city c at time t. Panel A reports 2SLS estimates, instrumenting for payment flow with bike placement; Panel B reports the first stage; Panel C reports OLS estimates. All columns use individual and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit Access$_{i,t}$		$log(\text{Credit Line})_{i,t}$	
	(1)	(2)	(3)	(4)
Panel A. Two-Stage Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.097***	0.098***	0.280***	0.282***
	(0.025)	(0.026)	(0.085)	(0.086)
$log(1 + \text{Assets under Management})_{i,t}$	-0.005	-0.008	-0.015	-0.026*
	(0.004)	(0.005)	(0.011)	(0.013)
Panel B. First Stage for $log(1 + \text{In-Person Payment Flow})_{i,t}$				
$log(\text{Bike Placement})_{c,t}$	0.038***	0.036***	0.043***	0.043***
	(0.010)	(0.010)	(0.011)	(0.011)
$log(1 + \text{Assets under Management})_{i,t}$	0.147***	0.180***	0.122***	0.152***
	(0.005)	(0.005)	(0.005)	(0.005)
F-Statistic	14.7	14.6	14.4	14.6
Adjusted R^2	0.562	0.566	0.533	0.536
Panel C. Ordinary Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.009***	0.008***	0.020***	0.020***
	(0.001)	(0.001)	(0.002)	(0.002)
$log(1 + \text{Assets under Management})_{i,t}$	0.008***	0.008***	0.017***	0.014***
	(0.001)	(0.001)	(0.002)	(0.002)
Adjusted R^2	0.741	0.742	0.836	0.836
Whether AUM Include Account Balance	NO	YES	NO	YES
Individual FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Has Credit	Has Credit
Observations	1,220,618	1,220,618	779,283	779,283
Note:				*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 8: Financially Underserved Segments

These table provides evidence that the less educated and the older tend to be financially underserved in China. Low Education$_i$ and Older than Median$_i$ are dummies for having no bachelor's degree or above and being older than half of the users. # Debit Cards$_i$ is the number of debit cards linked to Alipay in April 2021. $log(1 + \text{Max. AUM})_i$ is the log highest amount of AUM on Alipay from May 2017 to September 2020. # Investment Months$_i$ is the months since first using Alipay's wealth management service till April 2021. Pay with Real Name$_i$, Use Own Account$_i$, and Complete Profile$_i$ are dummies for passing real name verification, using own account, and completing profile information in Alipay as of April 2021. Regression results show that the less educated and the older have lower financial service usage and literacy. All columns use city and gender fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Financial Service Usage			Financial Literacy		
	# Debit Cards$_i$	$log(1 + \text{Max. AUM})_i$	# Investment Months$_i$	Pay with Real Name$_i$	Use Own Account$_i$	Complete Profile$_i$
	(1)	(2)	(3)	(4)	(5)	(6)
Low Education$_i$	-0.694***	-1.078***	-3.076***	-0.119***	-0.087***	-0.122***
	(0.046)	(0.075)	(0.282)	(0.006)	(0.008)	(0.008)
Older than Median$_i$	-0.863***	-0.671***	-2.512***	-0.191***	-0.223***	-0.089***
	(0.025)	(0.045)	(0.141)	(0.006)	(0.009)	(0.005)
Gender FE	YES	YES	YES	YES	YES	YES
City FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Clustered by City	YES	YES	YES	YES	YES	YES
Observations	39,459	39,459	39,459	39,459	39,459	39,459
Adjusted R^2	0.081	0.052	0.036	0.081	0.101	0.046
Note:						$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 9: Education, In-person Payment Flow, and Credit Provision

The table shows the causal relationship between in-person payment flow and BigTech credit for the less and more educated groups on both margins. Credit Access$_{i,t}$ is a dummy for having access to Alipay's virtual credit card at time t. log(Credit Line)$_{i,t}$ is the log credit line conditional on Credit Access$_{i,t}$. $log(1 +$ In-Person Payment Flow$)_{i,t}$ is the $log(1 + x)$ transformed amount of in-person payment flow through Alipay in CNY at time t. log(Bike Placement)$_{c,t}$) is the log number of active shared bikes in city c at time t. Panel A reports 2SLS estimates, instrumenting for payment flow with bike placement; Panel B reports the first stage; Panel C reports OLS estimates. All columns use individual and year-month fixed effects. Columns (1) and (3) use the less educated subsample with no college degree or above; columns (2) and (4) use the more educated subsample with a bachelor's degree or above. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit Access$_{i,t}$		log(Credit Line)$_{i,t}$	
	(1)	(2)	(3)	(4)
Panel A. Two-Stage Least Squares				
$log(1 +$ In-Person Payment Flow$)_{i,t}$	0.093***	0.024	0.334***	0.038
	(0.027)	(0.044)	(0.109)	(0.073)
Panel B. First Stage for $log(1 +$ In-Person Payment Flow$)_{i,t}$				
log(Bike Placement)$_{c,t}$	0.039***	0.043***	0.039***	0.053***
	(0.010)	(0.013)	(0.011)	(0.014)
F-Statistic	13.7	10.9	11.6	14.2
Adjusted R^2	0.554	0.563	0.528	0.483
Panel C. Ordinary Least Squares				
$log(1 +$ In-Person Payment Flow$)_{i,t}$	0.009***	0.013***	0.022***	0.013***
	(0.001)	(0.001)	(0.003)	(0.002)
Adjusted R^2	0.741	0.734	0.831	0.893
Individual FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Has Credit	Has Credit
Subsample	Low Education	High Education	Low Education	High Education
Observations	1,065,769	171,938	657,878	121,194

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 10: Counterfactual Analysis

These tables report the steady state comparisons between the simulated digital money user case with the simulated counterfactual, which is the new digital money adopter case. The only difference between the two cases is that in digital money user case, the lender knows borrowers' consumption, in addition to observed personal characteristics. Low Education$_i$ and Older than Median$_i$ are dummies for having no bachelor's degree or above and being older than half of the users. $\Delta log(\widehat{\text{Credit Line}_i}), \%$, $\Delta log(\widehat{\text{Consumer Welfare}_i}), \%$, $\Delta log(\widehat{\text{Lender Profit}_i}), \%$, and $\Delta log(\widehat{\text{Annualized Default Rate}_i}), \%$ are the percentage changes in credit line, consumer welfare, lender profit, and default rate from the new digital money adopter case to the digital money user case. Panel A reports results of the direct comparison of the steady states, showing how consumption revealed by payment data affects credit outcomes and welfare. Panel B reports the relative changes for groups with different personal characteristics. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

Panel A. Steady State Comparison: New Digital Money Adopter vs. Digital Money User

	Mean		Mean Difference	Relative Change
	New Digital Money Adopter	Digital Money User		
Credit Line$_i$	3,619.9	5,707.5	2,087.6	57.7%
Used Credit$_i$	1,562.6	1,780.4	217.7	13.9%
Consumer Welfare$_i$	1,209.9	1,222.5	12.6	1.0%
Lender Profit$_i$	45.3	50.6	5.2	11.6%
Total Welfare$_i$	1,255.2	1,273.0	17.8	1.4%
Annualized Default Rate$_i$	0.51%	0.58%	0.07%	13.3%

Panel B. Distributional Effects in the Steady State Comparison

	$\Delta log(\widehat{\text{Credit Line}_i}), \%$	$\Delta log(\widehat{\text{Consumer Welfare}_i}), \%$	$\Delta log(\widehat{\text{Lender Profit}_i}), \%$	$\Delta log(\widehat{\text{Annualized Default Rate}_i}), \%$
	(1)	(2)	(3)	(4)
Low Education$_i$	1.558**	0.036***	0.708***	0.007**
	(0.786)	(0.011)	(0.222)	(0.003)
Older than Median$_i$	1.164**	0.027***	0.392***	-0.001
	(0.530)	(0.007)	(0.150)	(0.002)
Male$_i$	1.326***	0.009	0.128	-0.0003
	(0.493)	(0.007)	(0.139)	(0.002)
City FE	YES	YES	YES	YES
Observations	38,008	38,008	38,008	38,008
R^2	0.031	0.006	0.009	0.007
Note:				*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

A Online Appendix

A.1 Econometric Framework

I use an econometric framework to clarify the economic environment and the assumptions for identification.

There are three parties in the economic environment: the BigTech company that provides both cashless payment services and consumer lending; the consumers who make decisions about making in-person purchases using cashless payment; and the bike-sharing company that makes decisions about when and where to place the shared bikes.

Since the BigTech company provides cashless payment services, it has access to payment flow information and can use it for credit evaluation. Thus, the BigTech credit line provided to a consumer is a function of the consumer's cashless payment flow. For tractability, the BigTech credit provision equation is assumed to take the following form:

$$cl_{i,t} = \alpha_0 + \alpha_1 \cdot ipf_{i,t} + \delta_i + \theta_t + \varepsilon_{i,t}^{OV} + \varepsilon_{i,t}^{EE} \tag{1}$$

where $cl_{i,t}$ is the credit line provided by the BigTech company to individual i at time t, $ipf_{i,t}$ is the in-person payment flow of individual i at time t, δ_i and θ_t are the individual-specific and time-specific characteristics that affect the credit provision, respectively, $\varepsilon_{i,t}^{OV}$ is the omitted variables that affect the credit line of individual i at time t, and $\varepsilon_{i,t}^{EE}$ is an exogenous error term that affects the credit line of individual i at time t.

For consumers, the decision to use in-person cashless payment depends not only on their personal characteristics and the time-specific shocks, but also the credit access provided to them by the BigTech company. With a higher credit line, the individual would have a more relaxed borrowing constraint while using the mobile wallet, which allows her to make a larger amount of cashless payments. Also, if an individual expects that she would get a higher credit line on the BigTech platform by using cashless payment more frequently, she might be encouraged to seek a higher BigTech credit line. For simplicity, the in-person cashless payment decision of individual i at time t is assumed to have a linear relationship with the credit line, and the corresponding equation is

$$ipf_{i,t} = \beta_0 + \beta_1 \cdot cl_{i,t} + \mu_i + \omega_t + \varphi_{i,t} \tag{2}$$

where μ_i and ω_t are the individual-specific and time-specific characteristics that affect the in-person payment flow decision, respectively. $\varphi_{i,t}$ is an exogenous error term that affects the in-person payment flow of individual i at time t.

For simplicity, individual-specific and time-specific characteristics are treated as vectors of dimension one. The parameter of interest to estimate is α_1 in the credit provision equation, which captures the direct effect of in-person payment flow on the credit line provided by the BigTech company. Since the BigTech credit provision and in-person cashless payment flow are jointly determined, there are simultaneity issues, and the ordinary least squares (OLS) estimate would be biased. Assuming that $\varepsilon^{EE}_{i,t} \perp \varphi_{i,t}$, the bias of the OLS estimate is captured in the following equation:

$$\begin{aligned}\hat{\alpha_1^{OLS}} &= \frac{Cov(cl_{i,t}, ipf_{i,t})}{Var(ipf_{i,t})} \\ &= \alpha_1 + \underbrace{\frac{1}{1-\alpha_1 \cdot \beta_1}}_{A} \cdot [\underbrace{\frac{Var(\delta_i + \theta_t + \varepsilon^{OV}_{i,t} + \varepsilon^{EE}_{i,t})}{Var(ipf_{i,t})} \cdot \beta_1}_{B} + \underbrace{\frac{Cov(\varepsilon^{OV}_{i,t}, \varphi_{i,t})}{Var(ipf_{i,t})}}_{C}]\end{aligned} \tag{3}$$

The bias is captured by $A \cdot (B + C)$, where $A = \frac{1}{1-\alpha_1 \cdot \beta_1}$, $B = \frac{Var(\delta_i+\theta_t+\varepsilon^{OV}_{i,t}+\varepsilon^{EE}_{i,t})}{Var(ipf_{i,t})} \cdot \beta_1$, $C = \frac{Cov(\varepsilon^{OV}_{i,t}, \varphi_{i,t})}{Var(ipf_{i,t})}$.

Assume that the bike placement decision $bp_{c,t}$ is exogenous and can be used a valid instrument for $ipf_{i,t}$. That is, $\mathsf{E}[(\varepsilon^{OV}_{i,t} + \varepsilon^{EE}_{i,t}) \cdot bp_{c,t}] = 0$ and $\mathsf{E}[\varphi_{i,t} \cdot bp_{c,t}] \neq 0$. The IV estimate is given by:

$$\hat{\alpha}_1^{IV} = \frac{Cov(cl_{i,t}, bp_{c,t})}{Cov(ipf_{i,t}, bp_{c,t})} = \alpha_1 \tag{4}$$

The econometric model does not provide direct predictions about how the magnitude of the IV estimate compares with the OLS estimate, but it helps to sort out the sources of the difference between the two estimates.

It is reasonable to assume that $0 < \alpha_1 < 1$ and $0 < \beta_1 < 1$, given the synergetic relationship between the cashless payment flow and the BigTech credit provision. With these assumptions, we get $A > 0$ and $B > 0$. The sign of C is determined by the covariance between the omitted variable term in the credit provision equation and the exogenous error term in the in-person cashless payment decision equation, $Cov(\varepsilon^{OV}_{i,t}, \varphi_{i,t})$. This term could either be positive or negative, depending on the types of the omitted variables. For example, if the omitted variable is a negative shock to the individual's health condition, its covariance with the shock in the in-person cashless payment equation should be negative, since the health shock is likely to increase spending on medicine and treatment and decrease the creditworthiness of the individual. On the other hand, if the omitted variable is a positive income shock, the covariance should be positive, since the income shock is likely to increase both the level of payment flow and the magnitude of credit provision.

A.2 An Illustrative Example on Effects of a Cashless Payment Shock

In the economy, there is a lender and a continuum of borrowers. The type of borrower i follows a uniform distribution between 0 and 1; that is, $\theta_i \sim U[0,1]$. Given the type of borrower θ_i, the lender chooses the optimal lending amount l_i to maximize its expected profit. If the lender decides not to lend, its profit is zero. When the lending amount is positive, there will be some uncertainties, and the expected profit will be creditworthiness-dependent. For example, the interest rate will be different for borrowers of different types, and the probability of repayment will depend on the creditworthiness, the lending amount, and the interest rate. To simplify the specification, I assume that the expected profit function takes the following form:

$$\pi_i(\theta_i, l_i) = \begin{cases} \theta_i + 2 \cdot \theta_i \cdot l_i - l_i^2 - 1 & \text{, if } l_i > 0 \\ 0 & \text{, if } l_i = 0 \end{cases} \tag{5}$$

This functional form has three properties. First, given the lending amount, the expected profit monotonically increases with the borrower creditworthiness. Second, there is an optimal lending amount, below which the expected profit increases with the lending amount and above which the expected profit decreases with the lending amount. Third, given the borrower's creditworthiness, if the optimal lending amount is nonzero, it strictly increases with the borrower's creditworthiness. With this specification in Eq. (5), the nonzero lending amount $l^*(\theta_i) = \theta_i$.

Three cases with different information provided to the lender are used to represent the stages the borrowers are cash users, the borrowers are new digital money adopters, and the borrowers are digital money users. The relationship between the optimal credit line and the creditworthiness of the borrower in different stages is illustrated in Figure 7. Figures (a) and (b) capture different scenarios with different thresholds in the digital money user stage.

In the first stage, or the cash user stage, borrower creditworthiness θ_i is fully unknown to the lender, which can only make the lending decision based on the distribution of borrower creditworthiness in the population. This captures the feature of the cash economy whereby transactions are not well recorded, and there is a lack of information about the creditworthiness of each borrower.

In the second stage, or the new digital money adopter stage, the lender receives a rough signal about the creditworthiness of borrower, which is specified as $s_i = \mathbb{1}(\theta_i \geq 0.25)$ in the financial divide scenario and as $s_i = \mathbb{1}(\theta_i \geq 0.8)$ in the financial inclusion scenario. This stage captures two facts when the borrowers are new digital money adopters. First, people will submit information about their personal characteristics, such as age, gender, and education when they register as new users of the digital wallet. Second, it is easier for wealthier individuals to prove their creditworthiness with the observed characteristics.

In the third stage, or the digital money user stage, the lender knows the exact creditworthiness of each borrower. This is a stage in which the digital payment system operated by the BigTech company covers almost all the customers and merchants, and the recorded cashless transactions render the information about the creditworthiness of everyone quite precise.

The lender makes very different credit-provision decisions in the stages with distinct information sets.

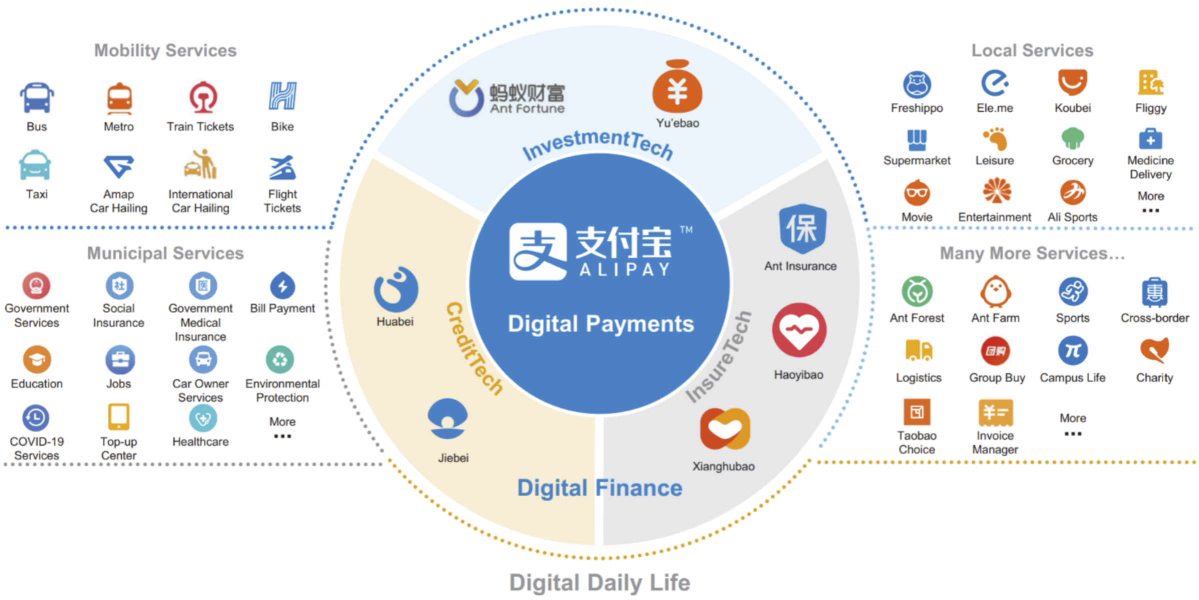
In the first stage, it knows only the distribution of the borrower creditworthiness and will make the same lending decision to every borrower based on the average creditworthiness of borrowers. Under above specification, lending a positive amount is always nonprofitable, and the lender will not lend to any borrower in this stage.

We first consider the financial divide scenario in figure (a), where the threshold in the second stage is 0.25. In the second stage, the lender knows whether each borrower i is the "high creditworthiness" with $\theta_i \geq 0.25$ or the "low creditworthiness" with $\theta_i < 0.25$. Intuitively, the lender will not lend to any low-creditworthiness borrower. For high-creditworthiness borrowers, it is optimal to lend $l^*(s_i = 1) = 0.625$ to everyone in this group, and this will maximize the expected profit from lending. It is not surprising that the rough signal helps the lender extend more credit in the transition from the first stage to the second stage. In the third stage, the lender has precise information on each borrower's creditworthiness, which enables it to make the optimal lending decision for each borrower creditworthiness separately. In this specification, the optimal lending decision is to not lend to borrowers with creditworthiness $\theta_i \leq \frac{\sqrt{5}-1}{2}$, and lend $l^*(\theta_i) = \theta_i$ to borrowers with $\theta_i > \frac{\sqrt{5}-1}{2}$. In the transition from the second stage to the third stage, although the borrowers with creditworthiness larger than 0.625 gain higher credit limit, the less creditworthy borrowers suffer from a large credit limit reduction, especially those with creditworthiness between 0.25 and $\frac{\sqrt{5}-1}{2}$. In this scenario, the gap of credit access between the less creditworthy and the more creditworthy becomes larger after the lender gets more precise information, and that is why we name it as "financial divide".

We then consider the financial inclusion scenario in figure (b), where the threshold in the second stage is 0.8. Now in the second stage, the lender knows whether each borrower i is the “high creditworthiness” with $\theta_i \geq 0.8$ or the “low creditworthiness” with $\theta_i < 0.8$. The lender will still not lend to any low-creditworthiness borrower. For high-creditworthiness borrowers, it is optimal to lend $l^*(s_i = 1) = 0.9$ to everyone in this group, and this will maximize the expected profit from lending. Comparing the second stage with the first stage, the rough signal helps the lender extend more credit. The third stage is the same as the financial divide scenario, while the comparison between the third stage and the second stage is different. In this transition, some of the previously underserved borrowers in the second stage $\frac{\sqrt{5}-1}{2} < \theta_i < 0.8$ now gain access to credit in the third stage. This is what we call ”financial inclusion”. Fore the high-creditworthiness borrowers, they get a creditworthiness-specific credit limit $l^*(\theta_i) = \theta_i$ instead of the same amount in the third stage, although the average lending amount stays at the level of 0.9.

Figure A.1: Typical Use Cases Available via the Alipay App

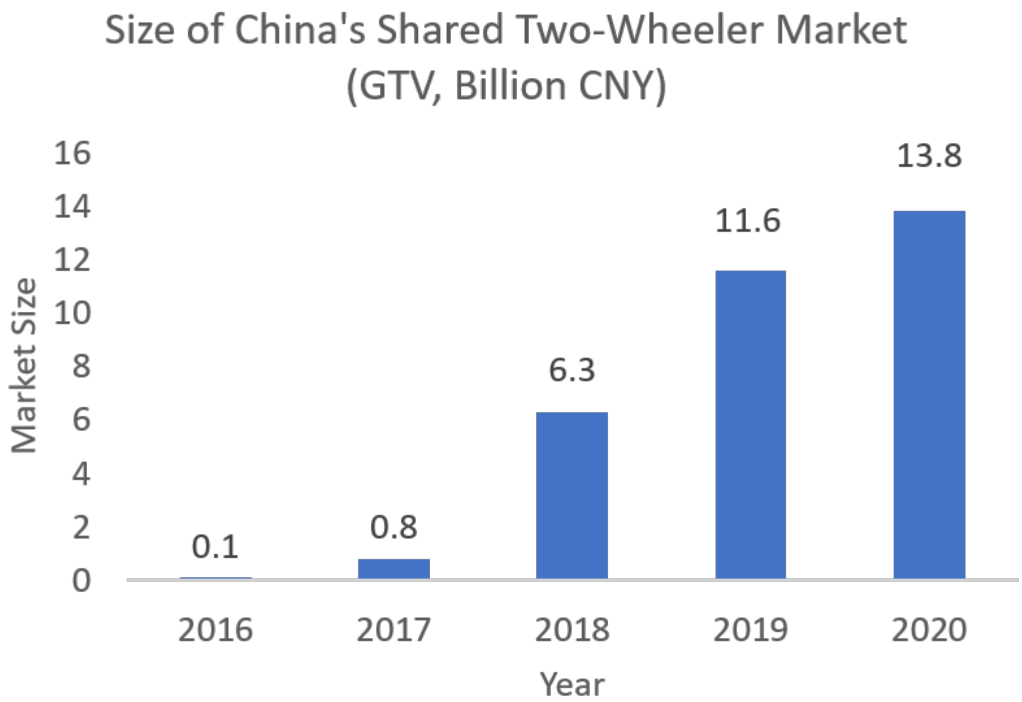
This figure describes the typical use cases that are available via the Alipay app, which cover mobility services, municipal services, local services, and other services. Alipay acts as consumers' one-stop shop for digital payment and digital financial services, including credit, investment, and insurance. There are over 1,000 daily life services and over 2 million mini-programs on Alipay.



Source: IPO Prospectus of the Ant Group, 2020

Figure A.2: Development of China's Dockless Bike-sharing Industry

This figure presents the time series of the size of China's shared two-wheeler market from 2016 to 2020. Market size is measured by the gross transaction volume (GTV) in billion CNY.



Source: IPO Prospectus of Hello Inc, 2021; iResearch Report

Figure A.3: Alipay Registration and Shared-bike Adoption

This bar plot presents the fraction of sampled users in four groups with different relationships between Alipay registration and bike adoption. *Adopt Bike in 1 Month* means that the user starts to use Alipay-bundled shared bikes within 1 month of registering with Alipay; *Adopt Bike in 2 to 12 Months* means that the user starts to use Alipay-bundled shared bikes more than 1 month but less than 1 year after registering with Alipay; *Adopt Bike Later than 1 Year* means that the user starts to use Alipay-bundled shared bikes more than 1 year after registering in Alipay; *Never Adopt Bike* means that the Alipay user has never used Alipay-bundled shared bikes during the sample period.

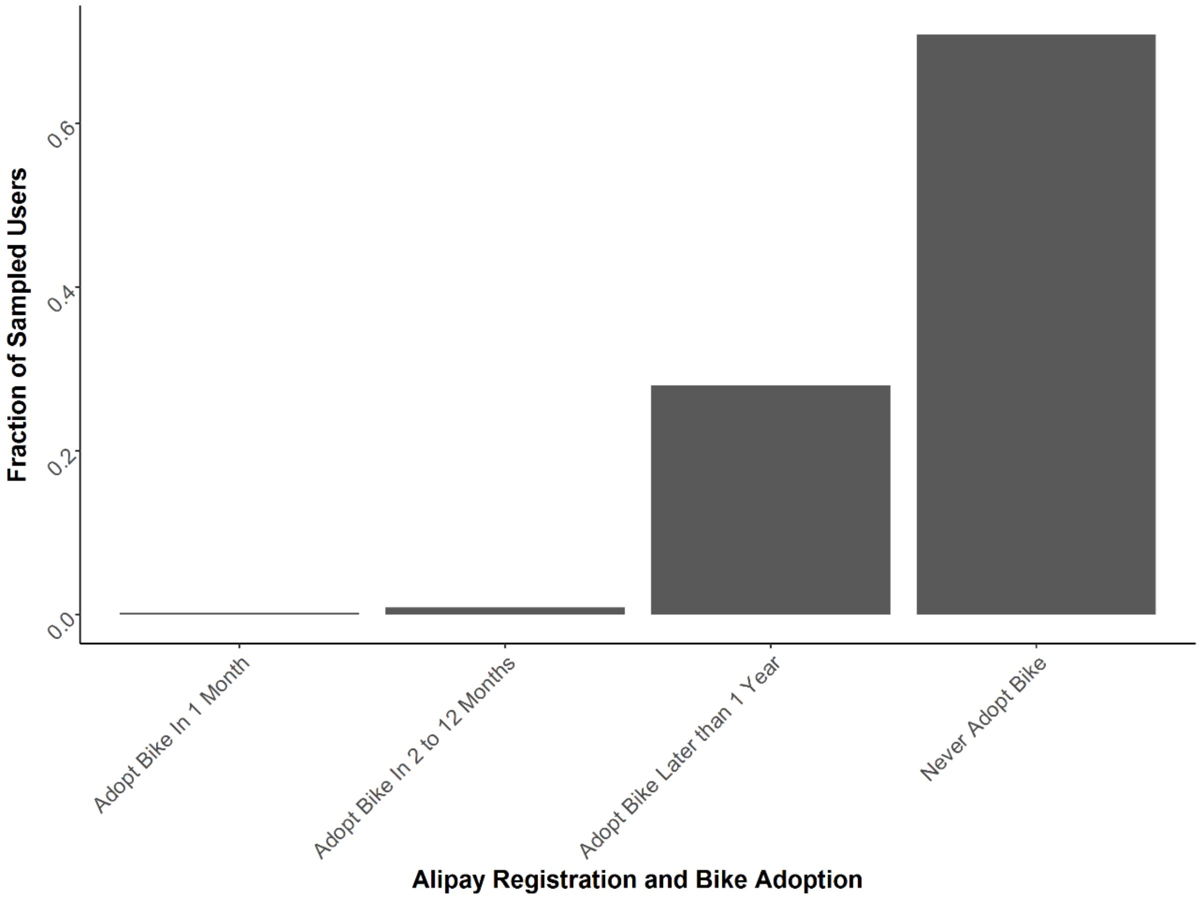


Figure A.4: Evidence of the Nonmonotone Payment-credit Relationship

This figure presents the fitted linear and quadratic relationship between the normalized credit line and the normalized in-person payment flow.

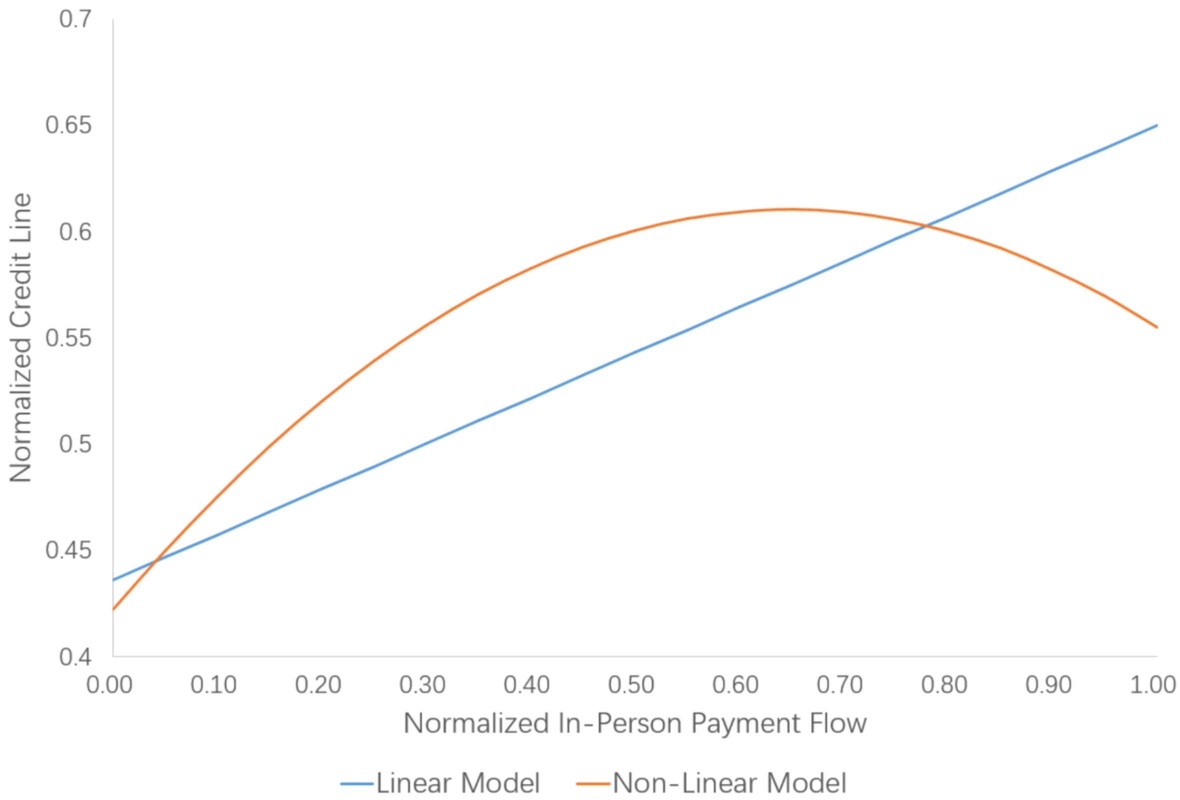


Figure A.5: Credit Access and Total Payment Flow

This figure plots the β_τ coefficients estimated in the following regression:

$$
\begin{aligned}
log(1 + \text{Total Payment Flow})_{i,t} &= \alpha_0 \\
&+ \sum_{\tau=-5}^{4} \beta_\tau \cdot \mathbb{1}(t=\tau) \cdot \mathbb{1}(\tau \neq -1) + \beta_5 \cdot \mathbb{1}(t \geq 5) + \delta_i + \mu_t + \varepsilon_{i,t}
\end{aligned}
$$

where $log(1 + \text{Total Payment Flow})_{i,t}$ is the $log(1+x)$ transformed amount of total payments by individual i at time t using Alipay, t corresponds to the number of months after each individual's first month of having access to Huabei credit line, δ_i is the individual fixed effects, μ_t is the year-month fixed effects, and $\varepsilon_{i,t}$ is the error term that varies across individuals and over time. The sample covers only users who have access to Huabei credit line in at least one month in the sample period, which is from May 2017 to September 2020. For each user, the sample only covers periods in which the t is not earlier than -5.

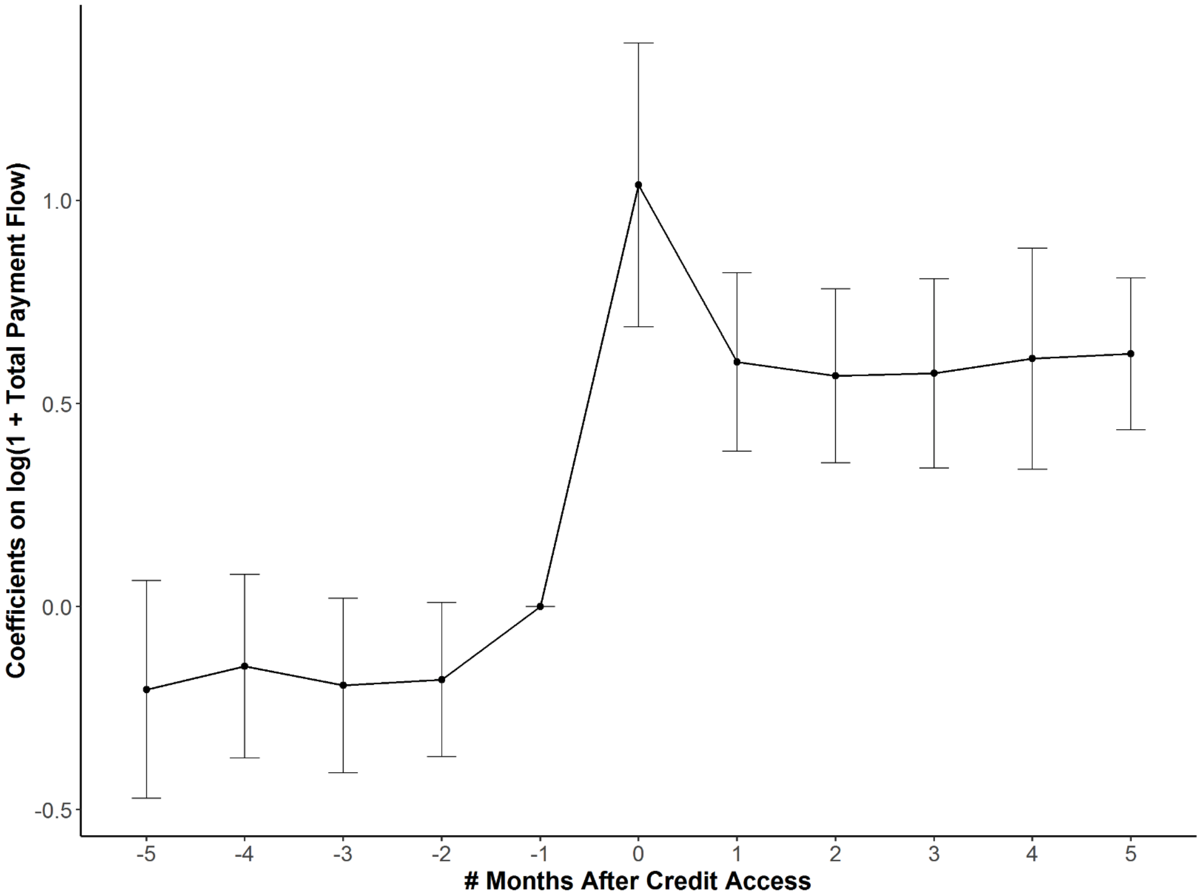


Table A.1: Effects of Bike Placement on Bike Usage

The table shows the effects of city-level bike placement on individual-level bike-riding activities. $log(\text{Bike Placement})_{c,t}$ is the log number of active shared bikes in city c at time t. Use $\text{Bike}_{i,t}$ is a dummy for using a shared bike at time t. $log(\text{\# Bike Rides})_{i,t}$ and $log(\text{Riding Distance})_{i,t}$ are the log number and distance of bike rides at time t. Column (1) uses the sample of bike users who have used a shared bike at least once from May 2017 to September 2020. Columns (2) and (3) use the sample of bike users in the months they used a bike. All columns use individual and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Use $\text{Bike}_{i,t}$	$log(\text{\# Bike Rides})_{i,t}$	$log(\text{Riding Distance})_{i,t}$
	(1)	(2)	(3)
$log(\text{Bike Placement})_{c,t}$	0.028***	0.102***	0.161***
	(0.003)	(0.014)	(0.040)
Individual FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES
Sample	Bike Users	Bike Users, Bike Using Months	Bike Users, Bike Using Months
Observations	435,872	69,978	66,048
Adjusted R^2	0.203	0.372	0.306
Note:			*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table A.2: Bike-riding Activity and Payment Flow

The table shows the relationship between bike-riding activity and cashless payment flow with and without bike-related spending. After First Bike Usage$_{i,t}$ is a dummy for after using a shared bike for the first time. $log(1 + \text{\# Bike Rides})_{i,t}$ and $log(1 + \text{Riding Distance})_{i,t}$ are the $log(1 + x)$ transformed number and distance of bike rides in kilometers at time t. $log(1 + \text{In-Person Payment Flow})_{i,t}$ and $log(1 + \text{In-Person Non-Bike Payment Flow})_{i,t}$ are the $log(1 + x)$ transformed amount of in-person payment flow in CNY through Alipay at time t, with and without bike-related spending. Columns (1) and (4) use the sample of users who have used shared bikes at least once and cover all their periods. Columns (2), (3), (5), and (6) use the same sample but only after they started using bikes. All columns use individual and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	$log(1 + \text{In-Person Payment Flow})_{i,t}$			$log(1 + \text{In-Person Non-Bike Payment Flow})_{i,t}$		
	(1)	(2)	(3)	(4)	(5)	(6)
After First Bike Usage$_{i,t}$	0.694***			0.638***		
	(0.055)			(0.053)		
$log(1 + \text{\# Bike Rides})_{i,t}$		0.347***			0.286***	
		(0.015)			(0.012)	
$log(1 + \text{Riding Distance})_{i,t}$			0.265***			0.211***
			(0.026)			(0.021)
Individual FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES	YES	YES
Sample	Bike Users	After First Ride	After First Ride	Bike Users	After First Ride	After First Ride
Observations	449,642	280,435	280,435	449,642	280,435	280,435
Adjusted R^2	0.484	0.528	0.527	0.483	0.526	0.525

Note: $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.3: Personal Characteristics of Bike Users

This table presents the relationship between an individual's personal characteristics and the bike user dummy, indicating whether the user has used Alipay-bundled shared bikes at least once. Low Education$_i$, Older than Median$_i$, Early Alipay User$_i$, Male$_i$, Pay with Real Name$_i$, Use Own Account$_i$, and Complete Profile$_i$ are dummy variables defined based on education, age, registration date, gender, real-name verification, account usage, and profile completion, respectively. Bike User$_i$ equals 1 if Alipay user i used shared bikes during May 2017 to September 2020. Columns (1), (2), and (3) show simple regression, regression with city and occupation fixed effects, and regression controlling for financial activity measures, respectively. Financial activity measures include # Debit Cards$_i$, $log(1 + \text{Max. AUM})_i$, and # Investment Months$_i$. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Bike User$_i$		
	(1)	(2)	(3)
Low Education$_i$	-0.173***	-0.109***	-0.065***
	(0.009)	(0.010)	(0.009)
Older than Median$_i$	-0.095***	-0.110***	-0.096***
	(0.005)	(0.005)	(0.004)
Early Alipay User$_i$	-0.129***	-0.113***	-0.030***
	(0.007)	(0.006)	(0.005)
Male$_i$	0.049***	0.059***	0.045***
	(0.004)	(0.004)	(0.004)
Pay with Real Name$_i$	0.088***	0.081***	0.012**
	(0.006)	(0.005)	(0.005)
Use Own Account$_i$	0.076***	0.071***	0.033***
	(0.006)	(0.005)	(0.005)
Complete Profile$_i$	0.012*	0.001	-0.012*
	(0.007)	(0.006)	(0.006)
Constant	0.421***		
	(0.013)		
City FE	NO	YES	YES
Occupation FE	NO	YES	YES
Controls Financial Activity Measures	NO	NO	YES
Clustered by City	YES	YES	YES
Observations	39,459	39,459	39,459
Adjusted R^2	0.123	0.178	0.260
Note:	$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$		

Table A.4: Analysis of the Heterogeneous Effects of Bike Placement

This table reports the heterogeneous effects of city-level placement of shared bikes on individual-level in-person payment flow and the digital credit provided to the user. $log(\text{Bike Placement})_{c,t}$ is the log-transformed count of active shared bikes in city c at time t. Bike User$i$, Low Education$_i$, Older than Median$_i$, Early Alipay User$_i$, Male$_i$, Pay with Real Name$_i$, and Use Own Account$_i$ are dummy variables based on bike usage, education, age, registration date, gender, real-name verification, and account usage, respectively. $log(1 + \text{In-Person Payment Flow})_{i,t}$ and $log(1 + \text{Credit Line})_{i,t}$ are log-transformed measures of individual i's in-person payment flow and virtual credit card credit line in CNY at time t. Panel A shows OLS regressions with $log(1 + \text{In-Person Payment Flow})_{i,t}$ as the dependent variable, while Panel B uses $log(1 + \text{Credit Line})_{i,t}$. The Characteristic Measure$_i$ varies by column. All regressions control for individual and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Dependent Variable					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Ordinary Least Squares with Dependent Variable: $log(1 + \text{In-Person Payment Flow})_{i,t}$						
$log(\text{Bike Placement})_{c,t}$	-0.022	0.008	0.029**	0.021**	-0.013	-0.010
	(0.014)	(0.010)	(0.011)	(0.009)	(0.015)	(0.010)
Bike User$_i$ × $log(\text{Bike Placement})_{c,t}$	0.139***	0.110***	0.092***	0.099***	0.057**	0.139***
	(0.029)	(0.018)	(0.017)	(0.021)	(0.025)	(0.029)
Characteristic Measure$_i$ × $log(\text{Bike Placement})_{c,t}$	0.036**	0.004	-0.038***	-0.023**	0.033*	0.036**
	(0.017)	(0.013)	(0.012)	(0.008)	(0.019)	(0.017)
Bike User$_i$ × Characteristic Measure$_i$ × $log(\text{Bike Placement})_{c,t}$	-0.040	-0.017	0.009	0.009	0.046**	-0.045
	(0.031)	(0.018)	(0.025)	(0.020)	(0.023)	(0.031)
Adjusted R^2	0.552	0.552	0.552	0.552	0.552	0.552
Panel B. Ordinary Least Squares with Dependent Variable: $log(1 + \text{Credit Line})_{i,t}$						
$log(\text{Bike Placement})_{c,t}$	0.009	0.014	0.020	0.004	-0.008	0.003
	(0.021)	(0.010)	(0.013)	(0.014)	(0.013)	(0.015)
Bike User$_i$ × $log(\text{Bike Placement})_{c,t}$	0.051*	0.053*	0.057*	0.056**	0.049*	0.042**
	(0.030)	(0.026)	(0.029)	(0.025)	(0.029)	(0.020)
Characteristic Measure$_i$ × $log(\text{Bike Placement})_{c,t}$	0.0001	-0.011	-0.023	0.008	0.024*	0.012
	(0.026)	(0.018)	(0.025)	(0.012)	(0.014)	(0.014)
Bike User$_i$ × Characteristic Measure$_i$ × $log(\text{Bike Placement})_{c,t}$	0.012	0.016	-0.008	0.007	0.007	0.022
	(0.025)	(0.028)	(0.046)	(0.019)	(0.037)	(0.034)
Adjusted R^2	0.800	0.799	0.800	0.799	0.800	0.800
Personal Characteristic Measure	Low Education$_i$	Older than Median$_i$	Early Alipay User$_i$	Male$_i$	Pay with Real Name$_i$	Use Own Account$_i$
Individual FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES	YES	YES
Observations	1,237,707	1,237,707	1,237,707	1,237,707	1,237,707	1,237,707

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.5: Bike Placement and the Local Economy

This table presents empirical evidence showing that conditional on city fixed effects and year-month fixed effects, city-level bike placement does not significantly correlate with key variables that describe local economic conditions. $log(\text{Bike Placement})_{c,t}$ is a log transformation of the number of active shared bikes placed in city c at time t. $log(\text{GDP})_{c,t}$ is the log of the gross domestic product (GDP) of city c at time t. $log(\text{GDP per capita})_{c,t}$ is the log of the GDP per capita in city c at time t. Fiscal Spending/GDP$_{c,t}$ is the ratio of local fiscal spending over the local GDP in city c at time t. Fiscal Income/GDP$_{c,t}$ is the ratio of local fiscal spending over the local GDP in city c at time t. All columns show results for the regressions with city fixed effects and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	$log(\text{GDP})_{c,t}$	$log(\text{GDP per capita})_{c,t}$	Fiscal Spending/GDP$_{c,t}$	Fiscal Income/GDP$_{c,t}$
	(1)	(2)	(3)	(4)
$log(\text{Bike Placement})_{c,t}$	0.002	0.000	-0.001	0.000
	(0.002)	(0.002)	(0.001)	(0.000)
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Clustered by City and Year	YES	YES	YES	YES
Observations	895	775	886	891
Adjusted R^2	0.992	0.979	0.957	0.903
Note:				$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.6: Nonmonotone Payment-Credit Relationship

This table reports the nonmonotone relationship between the normalized in-person payment flow and the normalized credit line. Normalized In-Person Payment Flow$_{i,t}$ is the total amount of individual i's in-person payment flow through Alipay at time t, normalized by their highest monthly in-person payment flow. Normalized Credit Line$_{i,t}$ is the credit line of Alipay user i's virtual credit card at time t, normalized by their highest credit line. Regressions in columns (1) and (2) are simple regressions without control variables, and regressions in columns (3) and (4) control for both individual fixed effects and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Normalized Credit Line$_{i,t}$			
	(1)	(2)	(3)	(4)
Normalized In-Person Payment Flow$_{i,t}$	0.214***	0.581***	0.040***	0.105***
	(0.033)	(0.076)	(0.006)	(0.013)
(Normalized In-Person Payment Flow$_{i,t}$)^2		-0.448***		-0.075***
		(0.064)		(0.009)
Constant	0.436***	0.422***		
	(0.042)	(0.043)		
Individual FE	NO	NO	YES	YES
Year-Month FE	NO	NO	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Observations	1,030,678	1,030,678	1,030,678	1,030,678
Adjusted R^2	0.016	0.022	0.767	0.767
Note:	*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$			

Table A.7: Robustness: In-person Payment Flow and Credit Provision, Controlling for City Times Year-Month Fixed Effects

This table presents empirical evidence of the causal link between a user's in-person payment flow and BigTech credit provided, controlling for city times year-month fixed effects on both extensive and intensive margins. Credit $\text{Access}_{i,t}$ is a binary variable indicating access to Alipay's virtual credit card for user i at time t. $log(\text{Credit Line})_i, t$ represents the log-transformed credit line of user i's virtual card at time t, missing if Credit $\text{Line}_{i,t}$ is 0. $log(1 + \text{In-Person Payment Flow})_{i,t}$ is the $log(1 + x)$ transformed in-person payment flow in CNY through Alipay for individual i at time t. $log(\text{Bike Placement})_{c,t}$ is the log-transformed count of active shared bikes in city c at time t. Bike User_i is 1 if user i used shared bikes at least once during May 2017-September 2020. Panel A reports 2SLS estimates, instrumenting individual-level log in-person payment flow with the interaction of individual-level bike user dummy and city-level log number of active shared bikes; Panel B reports the corresponding first stage. Panel C reports the OLS regression coefficient for individual-level log in-person payment flow. All columns show results for regressions with individual and year-month fixed effects. Columns (2) and (4) also control for individual characteristics like gender, education, occupation, and year of birth. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit $\text{Access}_{i,t}$		$log(\text{Credit Line})_{i,t}$	
	(1)	(2)	(3)	(4)
Panel A. Two-Stage Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.115***	0.108***	0.398***	0.418***
	(0.004)	(0.004)	(0.016)	(0.019)
Panel B. First Stage for $log(1 + \text{In-Person Payment Flow})_{i,t}$				
Bike $\text{User}_i \times log(\text{Bike Placement})_{c,t}$	0.209***	0.178***	0.166***	0.134***
	(0.008)	(0.008)	(0.007)	(0.007)
F-Statistic	772.9	476.0	503.2	343.0
Adjusted R^2	0.168	0.190	0.147	0.173
Panel C. Ordinary Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.054***	0.047***	0.147***	0.121***
	(0.001)	(0.001)	(0.004)	(0.004)
Adjusted R^2	0.193	0.245	0.181	0.363
City × Year-Month FE	YES	YES	YES	YES
Controls Individual Characteristics	NO	YES	NO	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Has Credit	Has Credit
Observations	1,238,309	664,727	779,283	440,418
Note:				*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table A.8: Robustness: In-person Payment Flow and Future Credit Provision

This table presents empirical evidence of the persistent relationship between a user's in-person payment flow and BigTech credit on both extensive and intensive margins. Credit Access$_{i,T}$ is a binary variable indicating access to Alipay's virtual credit card for user i at time T ($T = t+1, t+2, t+3$). $log(\text{Credit Line})_{i,T}$ represents the log-transformed credit line of user i's virtual card at time T, missing if Credit Line$_{i,T}$ is 0. $log(1 + \text{In-Person Payment Flow})_{i,t}$ is the $log(1+x)$ transformed in-person payment flow in CNY through Alipay for individual i at time t. $log(\text{Bike Placement})_{c,t}$ is the log-transformed count of active shared bikes in city c at time t. Panel A reports 2SLS estimates, instrumenting individual-level log in-person payment flow with city-level log number of active shared bikes; Panel B reports the corresponding first stage. Panel C reports the OLS regression coefficient for individual-level log in-person payment flow. All columns show results for regressions with individual and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit Access$_{i,T}$			$log(\text{Credit Line})_{i,T}$		
	$t+1$	$t+2$	$t+3$	$t+1$	$t+2$	$t+3$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Two-Stage Least Squares						
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.088***	0.085***	0.083***	0.250***	0.242***	0.235***
	(0.023)	(0.024)	(0.024)	(0.071)	(0.069)	(0.064)
Panel B. First Stage for $log(1 + \text{In-Person Payment Flow})_{i,t}$						
$log(\text{Bike Placement})_{c,t}$	0.041***	0.042***	0.042***	0.048***	0.048***	0.049***
	(0.011)	(0.011)	(0.011)	(0.012)	(0.013)	(0.013)
F-Statistic	15.4	15.1	15.4	15.0	14.6	15.0
Adjusted R^2	0.552	0.553	0.554	0.523	0.522	0.521
Panel C. Ordinary Least Squares						
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.008***	0.007***	0.006***	0.025***	0.026***	0.027***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Adjusted R^2	0.743	0.750	0.757	0.837	0.839	0.841
Individual FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Full Sample	Has Credit	Has Credit	Has Credit
Observations	1,199,746	1,161,435	1,123,295	775,512	763,560	750,694

Note: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table A.9: Robustness: In-person Payment Flow and Credit Provision, Controlling for Past Payment Flows

This table presents empirical evidence of the relationship between a user's in-person payment flow and BigTech credit, controlling for past payment flows on both extensive and intensive margins. Credit Access$_{i,t}$ is a binary variable indicating access to Alipay's virtual credit card for user i at time t. $log(\text{Credit Line})_{i,t}$ represents the log-transformed credit line of user i's virtual card at time t, missing if Credit Line$_{i,t}$ is 0. $log(1 + \text{In-Person Payment Flow})_{i,t}$ is the $log(1 + x)$ transformed in-person payment flow in CNY through Alipay for individual i at time t. $log(\text{Bike Placement})_{c,t}$ is the log-transformed count of active shared bikes in city c at time t. Panel A reports 2SLS estimates, instrumenting individual-level log in-person payment flow with city-level log number of active shared bikes; Panel B reports the corresponding first stage. Panel C reports the OLS regression coefficient for individual-level log in-person payment flow. Columns (1) and (4) control for past payment flow; columns (2) and (5) for past two periods; columns (3) and (6) for past three periods. All columns show results for regressions with individual and year-month fixed effects. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit Access$_{i,t}$			$log(\text{Credit Line})_{i,t}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Two-Stage Least Squares						
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.139***	0.154***	0.157***	0.388***	0.457***	0.531**
	(0.038)	(0.048)	(0.056)	(0.129)	(0.167)	(0.204)
Panel B. First Stage for $log(1 + \text{In-Person Payment Flow})_{i,t}$						
$log(\text{Bike Placement})_{c,t}$	0.024***	0.019***	0.016***	0.027***	0.022***	0.018***
	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)	(0.005)
F-Statistic	16.7	14.0	11.0	16.4	14.5	12.3
Adjusted R^2	0.636	0.647	0.651	0.596	0.605	0.608
Panel C. Ordinary Least Squares						
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.007***	0.006***	0.006***	0.015***	0.012***	0.010***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Adjusted R^2	0.743	0.751	0.759	0.837	0.840	0.842
Controls $log(1 + \text{In-Person Payment Flow})_{i,t-1}$	YES	YES	YES	YES	YES	YES
Controls $log(1 + \text{In-Person Payment Flow})_{i,t-2}$	NO	YES	YES	NO	YES	YES
Controls $log(1 + \text{In-Person Payment Flow})_{i,t-3}$	NO	NO	YES	NO	NO	YES
Individual FE	YES	YES	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Full Sample	Has Credit	Has Credit	Has Credit
Observations	1,199,825	1,161,573	1,123,548	775,601	763,711	750,940
Note:						$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.10: Robustness: In-person Payment Flow and Credit Provision, Controlling for Bike Usage

This table investigates the relationship between a user's in-person payment flow and BigTech credit on both margins, controlling for bike usage. Credit $\text{Access}_{i,t}$ is a dummy indicating access to Alipay's virtual credit card for user i at time t. $log(\text{Credit Line})_{i,t}$ represents the log-transformed credit line of user i's virtual card at time t conditional on positive Credit $\text{Line}_{i,t}$. $log(1 + \text{In-Person Payment Flow})_{i,t}$ is the $log(1 + x)$ transformed in-person payment flow in CNY through Alipay for individual i at time t. $log(\text{Bike Placement})_{c,t}$ is the log-transformed count of active shared bikes in city c at time t. Panel A displays 2SLS estimates, instrumenting individual-level log payment flow with city-level log active shared bikes. Panel B shows the first stage, and Panel C reports OLS regression coefficients. Regressions include individual and year-month fixed effects. All the standard errors are clustered at the city and year-month level. In columns (1) and (3), the measure of bike usage is the number of bike rides, while in columns (2) and (4) it is the riding distance measured in kilometers. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit $\text{Access}_{i,t}$		$log(\text{Credit Line})_{i,t}$	
	(1)	(2)	(3)	(4)
Panel A. Two-Stage Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.098***	0.097***	0.329***	0.329***
	(0.030)	(0.030)	(0.112)	(0.112)
$log(1 + \text{Measure of Bike Usage})_{i,t}$	-0.034**	-0.028**	-0.112**	-0.094**
	(0.015)	(0.012)	(0.048)	(0.041)
Panel B. First Stage for $log(1 + \text{In-Person Payment Flow})_{i,t}$				
$log(\text{Bike Placement})_{c,t}$	0.034***	0.034***	0.036***	0.036***
	(0.010)	(0.010)	(0.011)	(0.011)
$log(1 + \text{Measure of Bike Usage})_{i,t}$	0.497***	0.391***	0.408***	0.324***
	(0.022)	(0.030)	(0.021)	(0.027)
F-Statistic	11.2	11.2	10.2	10.2
Adjusted R^2	0.554	0.554	0.530	0.529
Panel C. Ordinary Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.010***	0.010***	0.021***	0.022***
	(0.001)	(0.001)	(0.003)	(0.003)
$log(1 + \text{Measure of Bike Usage})_{i,t}$	0.010***	0.007***	0.015***	0.007*
	(0.002)	(0.001)	(0.005)	(0.004)
Adjusted R^2	0.740	0.740	0.836	0.836
Measure of Bike Usage	# Bike Rides	Riding Distance	# Bike Rides	Riding Distance
Individual FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Has Credit	Has Credit
Observations	1,238,309	1,238,309	779,283	779,283
Note:				$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.11: Robustness: In-person Payment Flow and Credit Provision, Controlling for Online Payment

This table investigates the relationship between a user's in-person payment flow and BigTech credit on both margins, controlling for online payment. Credit Access$_{i,t}$ is a dummy indicating access to Alipay's virtual credit card for user i at time t. $log(\text{Credit Line})_{i,t}$ represents the log-transformed credit line of user i's virtual card at time t conditional on positive Credit Line$_{i,t}$. $log(1 + \text{In-Person Payment Flow})_{i,t}$ is the $log(1 + x)$ transformed in-person payment flow in CNY through Alipay for individual i at time t. $log(\text{Bike Placement})_{c,t}$ is the log-transformed count of active shared bikes in city c at time t. Panel A displays 2SLS estimates, instrumenting individual-level log payment flow with city-level log active shared bikes. Panel B shows the first stage, and Panel C reports OLS regression coefficients. Regressions include individual and year-month fixed effects. All the standard errors are clustered at the city and year-month level. In columns (1) and (3), the measure of online payment is the online payment flow measured in CNY, while in columns (2) and (4) it is the number of online transactions. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit Access$_{i,t}$		$log(\text{Credit Line})_{i,t}$	
	(1)	(2)	(3)	(4)
Panel A. Two-Stage Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.086***	0.085***	0.280***	0.277***
	(0.023)	(0.023)	(0.085)	(0.082)
$log(1 + \text{Measure of Online Payment})_{i,t}$	-0.009	-0.028	-0.037*	-0.107*
	(0.006)	(0.017)	(0.021)	(0.054)
Panel B. First Stage for $log(1 + \text{In-Person Payment Flow})_{i,t}$				
$log(\text{Bike Placement})_{c,t}$	0.041***	0.042***	0.043***	0.044***
	(0.010)	(0.010)	(0.012)	(0.012)
$log(1 + \text{Measure of Online Payment})_{i,t}$	0.260***	0.716***	0.246***	0.649***
	(0.007)	(0.015)	(0.008)	(0.018)
F-Statistic	16.0	16.2	14.0	14.3
Adjusted R^2	0.572	0.574	0.544	0.545
Panel C. Ordinary Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.008***	0.008***	0.018***	0.018***
	(0.001)	(0.001)	(0.002)	(0.002)
$log(1 + \text{Measure of Online Payment})_{i,t}$	0.011***	0.027***	0.027***	0.061***
	(0.001)	(0.002)	(0.003)	(0.007)
Adjusted R^2	0.742	0.742	0.837	0.836
Measure of Online Payment	Online Payment Flow	# Online Transactions	Online Payment Flow	# Online Transactions
Individual FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Has Credit	Has Credit
Observations	1,238,309	1,238,309	779,283	779,283

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.12: Age, In-person Payment Flow, and Credit Provision

This table provides empirical evidence of the causal link between a user's in-person payment flow and BigTech credit for older and younger groups on both extensive and intensive margins. Credit $\text{Access}_{i,t}$ is a binary variable for Alipay user i's virtual credit card access at time t. $log(\text{Credit Line})_{i,t}$ denotes the log-transformed credit line for user i at t. $log(1 + \text{In-Person Payment Flow})_{i,t}$ represents the $log(1 + x)$ transformed payment flow through Alipay for user i at t, measured in CNY. $log(\text{Bike Placement})_{c,t}$ signifies the log-transformed active shared bikes in city c at t. Panel A displays 2SLS estimates, instrumenting individual-level log payment flow with city-level log active shared bikes; Panel B shows the first stage. Panel C reports OLS regression coefficients against individual-level log payment flow. Regressions include individual and year-month fixed effects. Columns (1) and (3) use the subsample of older people, who are older than more than half of the individuals in the sample; columns (2) and (4) use the subsample of younger people, who are not older than half of the individuals in the sample. All standard errors are clustered at city and year-month level. ***, **, and * denote the 1%, 5%, and 10% confidence level, respectively. I report standard errors in parentheses.

	Credit $\text{Access}_{i,t}$		$log(\text{Credit Line})_{i,t}$	
	(1)	(2)	(3)	(4)
Panel A. Two-Stage Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.124***	0.047**	0.440***	0.176**
	(0.041)	(0.020)	(0.177)	(0.065)
Panel B. First Stage for $log(1 + \text{In-Person Payment Flow})_{i,t}$				
$log(\text{Bike Placement})_{c,t}$	0.032***	0.049***	0.030***	0.054***
	(0.010)	(0.012)	(0.011)	(0.013)
F-Statistic	9.7	17.8	7.0	16.6
Adjusted R^2	0.552	0.539	0.559	0.483
Panel C. Ordinary Least Squares				
$log(1 + \text{In-Person Payment Flow})_{i,t}$	0.009***	0.011***	0.017***	0.026***
	(0.001)	(0.001)	(0.003)	(0.002)
Adjusted R^2	0.739	0.740	0.833	0.847
Individual FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Clustered by City and Year-Month	YES	YES	YES	YES
Sample	Full Sample	Full Sample	Has Credit	Has Credit
Subsample	Older than Median	Younger than Median	Older than Median	Younger than Median
Observations	577,711	654,823	335,670	443,402

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.13: Distribution of Observables in the Structural Estimation

This table presents the summary statistics of the key observables in the structural estimation. c_i is the consumption of borrower i. b_i is the credit line used by borrower i. e_{i1} is the realized income to borrower i in period $t = 1$. l_i is the credit limit provided by the lender to borrower i.

	N	Mean	Std	Min	p25	Median	p75	Max
c_i	38,276	1,595.1	2,049.9	0.0	134.4	715.5	2,210.5	7,606.7
b_i	38,276	487.7	732.9	0.0	0.0	56.3	731.0	2,377.8
e_{i1}	38,276	1,122.0	1,665.8	0.0	48.8	344.2	1,431.9	7,606.7
l_i	38,276	7,145.5	10,256.8	0.0	0.0	3,000.0	10,000.0	61,000.0

References

Agarwal, S., Alok, S., Ghosh, P., and Gupta, S. (2021a). Financial Inclusion and Alternate Credit Scoring: Role of Big Data and Machine Learning in Fintech. SSRN Scholarly Paper ID 3507827, Social Science Research Network, Rochester, NY.

Agarwal, S., Kigabo, T., Minoiu, C., Presbitero, A., and Silva, A. F. (2021b). Serving the Underserved: Microcredit as a Pathway to Commercial Banks.

Ausubel, L. M. (1991). The Failure of Competition in the Credit Card Market. *The American Economic Review*, 81(1):50–81. Publisher: American Economic Association.

Babina, T., Buchak, G., and Gornall, W. (2022). Customer Data Access and Fintech Entry: Early Evidence from Open Banking. SSRN Scholarly Paper 4071214, Social Science Research Network, Rochester, NY.

Bachas, P., Gertler, P., Higgins, S., and Seira, E. (2021). How Debit Cards Enable the Poor to Save More. *The Journal of Finance*, 76(4):1913–1957. Publisher: John Wiley & Sons, Ltd.

Beaumont, P., Tang, H., and Vansteenberghe, E. (2022). The Role of FinTech in Small Business Lending. page 52.

Beck, T., Gambacorta, L., Huang, Y., Li, Z., and Qiu, H. (2022). Big Techs, QR Code Payments and Financial Inclusion.

Begley, T. A. and Purnanandam, A. (2021). Color and Credit: Race, Regulation, and the Quality of Financial Services. *Journal of Financial Economics*, 141(1):48–65.

Benetton, M., Buchak, G., and Robles-Garcia, C. (2022). Term's Up; Banking on Credit Cards.

Berg, T., Burg, V., Gombović, A., and Puri, M. (2020). On the Rise of FinTechs: Credit Scoring Using Digital Footprints. *The Review of Financial Studies*, 33(7):2845–2897.

Berg, T., Fuster, A., and Puri, M. (2021). FinTech Lending. Working Paper 29421, National Bureau of Economic Research. Series: Working Paper Series.

Bjorkegren, D., Blumenstock, J. E., and Knight, S. (2020). Manipulation-Proof Machine Learning. page 53.

Black, F. (1975). Bank Funds Management in an Efficient Market. *Journal of Financial Economics*, 2(4):323–339.

Brown, J. R., Cookson, J. A., and Heimer, R. Z. (2019). Growing Up Without Finance. *Journal of Financial Economics*, 134(3):591–616.

Brown, M., Ongena, S., Popov, A., and Yeşin, P. (2011). Who Needs Credit and Who Gets Credit in Eastern Europe? *Economic Policy*, 26(65):93–130.

Brunnermeier, M. and Payne, J. (2022). Platforms, Tokens, and Interoperability. page 42.

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal Expectations. *American Economic Review*, 95(4):1092–1118.

Buchak, G., Matvos, G., Piskorski, T., and Seru, A. (2018). Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks. *Journal of Financial Economics*, 130(3):453–483.

Cao, G., Jin, G. Z., Weng, X., and Zhou, L.-A. (2018). Market Expanding or Market Stealing? Competition with Network Effects in Bike-Sharing. *NBER Working Paper*. 1.

Chatterjee, S., Corbae, D., Dempsey, K. P., and Ríos-Rull, J.-V. (2020). A Quantitative Theory of the Credit Score.

Chen, L., Bolton, P., Holmström, B. R., Maskin, E., Pissarides, C. A., Spence, A. M., Sun, T., Sun, T., Xiong, W., Yang, L., Huang, Y., Li, Y., Luo, X., Ma, Y., Ouyang, S., and Zhu, F. (2021a). Understanding Big Data: Data Calculus in the Digital Era. SSRN Scholarly Paper ID 3791018, Social Science Research Network, Rochester, NY.

Chen, L., Huang, Y., Ouyang, S., and Xiong, W. (2021b). Data Privacy and Digital Demand.

Chen, T., Huang, Y., Lin, C., and Sheng, Z. (2021c). Finance and Firm Volatility: Evidence from Small Business Lending in China. *Management Science*. Publisher: INFORMS.

Chopra, Y., Prabhala, N., and Tantri, P. L. (2017). Bank Accounts for the Unbanked: Evidence from a Big Bang Experiment.

Cong, L. W. and Mayer, S. (2022). The Coming Battle of Digital Currencies. page 64.

Cornelli, G., Frost, J., Gambacorta, L., Rau, R., Wardrop, R., and Ziegler, T. (2020). Fintech and Big Tech Credit: A New Database.

Cornelli, G., Frost, J., Gambacorta, L., Rau, R., Wardrop, R., and Ziegler, T. (2021). Fintech and Big Tech Credit: What Explains the Rise of Digital Lending? *RESEARCH REPORT*, page 5.

Célerier, C. and Matray, A. (2019). Bank-Branch Supply, Financial Inclusion, and Wealth Accumulation. *The Review of Financial Studies*, 32(12):4767–4809.

DeFusco, A. A., Tang, H., and Yannelis, C. (2021). Measuring the Welfare Cost of Asymmetric Information in Consumer Credit Markets. Working Paper 29270, National Bureau of Economic Research. Series: Working Paper Series.

deHaan, E., Kim, J., Lourie, B., and Zhu, C. (2022). Buy Now Pay (Pain?) Later.

Di Maggio, M., Katz, J., and Williams, E. (2022). Buy Now, Pay Later Credit: User Characteristics and Effects on Spending Patterns.

Di Maggio, M., Ratnadiwakara, D., and Carmichael, D. (2021). Invisible Primes: Fintech Lending with Alternative Data. SSRN Scholarly Paper ID 3937438, Social Science Research Network, Rochester, NY.

Di Maggio, M. and Yao, V. (2020). Fintech Borrowers: Lax Screening or Cream-Skimming? *The Review of Financial Studies*, (hhaa142).

Doornik, B. F. N. V., Farboodi, M., Schoenherr, D., and Skrastins, J. (2021). Tell Me Your Type: Information Asymmetry and Credit Allocation Through Consorcios. *Working Paper*.

Economides, N. and Jeziorski, P. (2017). Mobile Money in Tanzania. *Marketing Science*, 36(6):815–837. 1.

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

Fama, E. F. (1985). What's Different About Banks? *Journal of Monetary Economics*, 15(1):29–39.

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., and Zbinden, P. (2019). BigTech and the Changing Structure of Financial Intermediation. *Economic Policy*, 34(100):761–799.

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

Fuster, A., Plosser, M., Schnabl, P., and Vickery, J. (2019). The Role of Technology in Mortgage Lending. *The Review of Financial Studies*, 32(5):1854–1899.

Gambacorta, L., Huang, Y., Li, Z., Qiu, H., and Chen, S. (2022). Data versus Collateral. *Review of Finance*, page rfac022.

Ghosh, P., Vallee, B., and Zeng, Y. (2022). FinTech Lending and Cashless Payments. *SSRN Electronic Journal*.

Goldstein, I., Huang, C., and Yang, L. (2022). Open Banking under Maturity Transformation. *Working Paper*, page 53.

Han, L., Fraser, S., and Storey, D. J. (2009). Are Good or Bad Borrowers Discouraged from Applying for Loans? Evidence from Us Small Business Credit Markets. *Journal of Banking & Finance*, 33(2):415–424.

Han, P. and Wang, Z. (2021). Technology Adoption and Leapfrogging: Racing for Mobile Payments. page 44. 1.

Hau, H., Huang, Y., Shan, H., and Sheng, Z. (2019). How FinTech Enters China's Credit Market. *AEA Papers and Proceedings*, 109:60–64.

Hau, H., Huang, Y., Shan, H., and Sheng, Z. (2021). FinTech Credit and Entrepreneurial Growth.

He, Z., Huang, J., and Zhou, J. (2022). Open Banking: Credit Market Competition When Borrowers Own the Data.

Helms, B., Poorest, C. G. t. A. t., and Bank, W. (2006). *Access for All: Building Inclusive Financial Systems*. World Bank. Google-Books-ID: owm3AAAAIAAJ.

Hong, C. Y., Lu, X., and Pan, J. (2020). FinTech Adoption and Household Risk-Taking. *Working Paper*, page 60. 1.

Howell, S. T., Kuchler, T., Snitkof, D., Stroebel, J., and Wong, J. (2021). Lender Automation and Racial Disparities in Credit Access.

Huang, J. (2022). Fintech Expansion. *Working Paper*, page 66.

Jack, W. and Suri, T. (2014). Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution. *American Economic Review*, 104(1):183–223. 1.

Jansen, M., Nagel, F., Yannelis, C., and Zhang, A. L. (2022). Data and Welfare in Credit Markets. *SSRN Electronic Journal*.

Karlan, D. and Zinman, J. (2010). Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts. *The Review of Financial Studies*, 23(1):433–464.

Kiyotaki, N. and Moore, J. (1997). Credit Cycles. *Journal of Political Economy*, 105(2):211–248.

Laibson, D., Repetto, A., and Tobacman, J. (2000). A Debt Puzzle. Technical Report w7879, National Bureau of Economic Research.

Lee, D. S., McCrary, J., Moreira, M. J., and Porter, J. R. (2021). Valid t-ratio Inference for IV.

Liberti, J. M. and Petersen, M. A. (2019). Information: Hard and Soft. *The Review of Corporate Finance Studies*, 8(1):1–41.

Liu, L., Lu, G., and Xiong, W. (2022). The Big Tech Lending Model.

Lyons, A. C., Grable, J., and Zeng, T. (2019). Impacts of Financial Literacy on the Loan Decisions of Financially Excluded Households in the People's Republic of China. *SSRN Electronic Journal*.

Mbiti, I. and Weil, D. N. (2015). Mobile Banking: The Impact of M-Pesa in Kenya. In *NBER Chapters*, pages 247–293. National Bureau of Economic Research, Inc.

Medina, P. C. (2021). Side Effects of Nudging: Evidence from a Randomized Intervention in the Credit Card Market. *The Review of Financial Studies*, 34(5):2580–2607.

Melzer, B. T. (2011). The Real Costs of Credit Access: Evidence from the Payday Lending Market. *The Quarterly Journal of Economics*, 126(1):517–555.

Mian, A. and Sufi, A. (2011). House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis. *The American Economic Review*, 101(5):2132–2156. Publisher: American Economic Association.

Morse, A. (2011). Payday Lenders: Heroes or Villains? *Journal of Financial Economics*, 102(1):28–44.

Ordoñez, G., Perez-Reyna, D., and Yogo, M. (2019). Leverage Dynamics and Credit Quality. *Journal of Economic Theory*, 183:183–212.

Ouma, S. A., Odongo, T. M., and Were, M. (2017). Mobile Financial Services and Financial Inclusion: Is It a Boon for Savings Mobilization? *Review of Development Finance*, 7(1):29–35.

Ouyang, S. and Peng, C. (2022). Consumer Demand for Digital Money.

Parlour, C. A., Rajan, U., and Zhu, H. (2022). When FinTech Competes for Payment Flows. *The Review of Financial Studies*, 35(11):4985–5024.

Rampini, A. A. and Viswanathan, S. (2020). Collateral and Secured Debt. page 65.

Riley, E. (2018). Mobile Money and Risk Sharing Against Village Shocks. *Journal of Development Economics*, 135:43–58. 1.

Rishabh, K. (2022). Can Open Banking Substitute Credit Bureaus? page 41.

Stango, V. and Zinman, J. (2014). Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees. *The Review of Financial Studies*, 27(4):990–1030.

Stein, L. C. D. and Yannelis, C. (2020). Financial Inclusion, Human Capital, and Wealth Accumulation: Evidence from the Freedman's Savings Bank. *The Review of Financial Studies*, 33(11):5333–5377.

Stock, J. H. and Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. page 29.

Suri, T. and Jack, W. (2016). The Long-Run Poverty and Gender Impacts of Mobile Money. *Science*, 354(6317):1288–1292. 1.

Tang, H. (2019). Peer-to-Peer Lenders Versus Banks: Substitutes or Complements? *The Review of Financial Studies*, 32(5):1900–1938.

Tang, L., Xuan, Y., and Zou, H. (2022). Does Mobile Payment Adoption Reduce Corruption? Cross-country Evidence.

Zinman, J. (2015). Household Debt: Facts, Puzzles, Theories, and Policies. *Annual Review of Economics*, 7(1):251–276. _eprint: https://doi.org/10.1146/annurev-economics-080614-115640.