Review of Finance, 2023, 369–398
https://doi.org/10.1093/rof/rfac022
Advance Access Publication Date: 15 April 2022

Data versus Collateral*

Leonardo Gambacorta[1,2], Yiping Huang[3], Zhenhua Li[4], Han Qiu[1], and Shu Chen[4]

[1]Bank for International Settlements, Switzerland, [2]CEPR, UK, [3]Institute of Digital Finance and National School of Development, Peking University, China, and [4]Ant Group, China

Abstract

Using a unique dataset of more than 2 million Chinese firms that received credit from both an important big tech firm (Ant Group) and traditional commercial banks, this paper investigates how different forms of credit correlate with local economic activity, house prices, and firm characteristics. We find that big tech credit does not correlate with local business conditions and house prices when controlling for demand factors, but reacts strongly to changes in firm characteristics, such as transaction volumes and network scores used to calculate firm credit ratings. By contrast, both secured and unsecured bank credit react significantly to local house prices, which incorporate useful information on the environment in which clients operate and on their creditworthiness. This evidence implies that the wider use of big tech credit could reduce the importance of the collateral channel but, at the same time, make lending more reactive to changes in firms' business activity.

JEL classification: D22, G31, R30

* We thank Zhigou He, Yongxiang Wang, and two anonymous referees for very helpful comments and suggestions. We also thank Claudio Borio, Gong Cheng, Sebastian Doerr, Torsten Ehlers, Jon Frost, Wei Jiang, Catherine Koch, Jun Pan, Amit Seru, Hyun Song Shin, Michael Song and conference participants at ABFER, BIS, Banca d'Italia, Banco de Portugal, Bangor University, Banque de France, Catholic University Milan, ECB, European Finance Association, IADI, IMF, and Moscow University for useful comments and insights. Finally, we thank Giluio Cornelli, Guangyao Zhu, Fang Wang, Yongguo Li, Zhiyun Chen, Jinyan Huang, Yiteng Zhai, Yanming Fang, Xin Li, Zhengjun Nie, Liang Guo, Ting Xu, Peng Liu, Ma Li, and Ruiping Wu for help with the data, comments, and suggestions. The views in this paper are those of the authors only and do not necessarily reflect those of the Bank for International Settlements or Ant Group. We highlight that the data and analysis reported in this paper may contain errors and are not suited for the purpose of company valuation or to deduce conclusions about the business success and/or commercial strategy of Ant Group other firms. All statements made reflect the private opinions of the authors and do not express any official position of Ant Group and its management. The analysis was undertaken in strict observance of the Chinese law on privacy. Y.H. gratefully acknowledge financial support by the National Social Science Foundation of China (project number 18ZDA091). The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. Z.L. and S.C. disclose an employment relationship in Ant Group. Ant Group did not exercise any influence on the content of this paper, but has ensured confidentiality of the (raw) data.

Keywords: Big tech, Big data, Collateral, Banks, Asymmetric information, Credit markets

Received October 11, 2020; accepted March 21, 2022 by Editor Zhiguo He.

1. Introduction

Collateral is used in debt contracts to mitigate agency problems arising from asymmetric information. Banks usually require their borrowers to pledge tangible assets, such as real estate, to lessen ex-ante adverse selection problems (Bester, 1985; Chan and Kanatas, 1985; Besanko and Thakor, 1987) or as a way to reduce ex-post frictions, such as moral hazard (Aghion and Bolton, 1997; Holmström and Tirole, 1997), costly state verification (Gale and Hellwig, 1985; Boyd and Smith 1994; Cooley, Marimon, and Quadrini, 2004), and imperfect contract enforcement (Albuquerque and Hopenhayn, 2004).[1]

The use of collateral is more widespread for opaque firms, such as small- and medium-sized enterprises (SMEs) all over the world (FSB, 2019). The percentage of bank loans to SMEs that are collateralized is 53% in China (OECD, 2019), where many firms lack basic documentation and are geographically remote from bank branches.

With the development of fintech, especially the entry of large technology firms (big techs) into the provision of financial services, non-traditional data play an increasingly important role in credit assessment for SMEs (BIS, 2019). The use of big data and machine-learning techniques could help to reduce the importance of collateral in solving asymmetric information problems in credit markets. Using a unique dataset of more than 2 million Chinese firms that received credit from both an important big tech firm (Ant Group) and traditional commercial banks, this paper investigates how big tech and bank credit correlates with local business conditions, house prices, and firm characteristics.

The business model of big techs rests on enabling direct interactions among a large number of users on digital platforms. An essential by-product is their large stock of user data, which they use as an input to offer a range of services that exploit natural network effects, generating further user activity. Increased user activity then completes the circle, as it generates yet more data. The mutually reinforcing data-network-activity (DNA) feedback loop helps big tech firms to identify the characteristics of their clients and offer them financial services that best suit their needs. As a result, big techs can have a competitive advantage over banks and serve firms that otherwise would remain unbanked. Recent work suggests that big techs' credit scoring applied to small vendors outperforms models based on credit bureau ratings and traditional borrower characteristics (Frost *et al.*, 2019). All this could help to significantly advance financial inclusion and improve firms' performance (see Luohan Academy, 2019; Hau *et al.*, 2021).

By leveraging artificial intelligence, big techs could address asymmetric information problems differently from banks. Big tech and fintech credit may use alternative data sources, including insights gained from social media activity (Jagtiani and Lemieux, 2018a) and users' digital footprints (Berg *et al.*, 2020). In this new way of conducting financial intermediation, data could take the place of collateral.[2]

1 For a review of the literature on the effects of collateral in credit markets see, among others, Ioannidou, Pavanini, and Peng (2022).

2 This is the reflection of the general principle that financial intermediaries substitute information for collateral when collateral is relatively more expensive (Holmström and Tirole, 1997).

Access to big data is not the only potential advantage for big techs over banks. Big techs have the further advantage of being able to monitor borrowers once they are within a big tech's ecosystem. For example, when a borrower is closely integrated into an e-commerce platform, it may be relatively easy for a big tech to deduct the (monthly) payments on a credit line from the borrower's revenues that pass through its payment account. This is useful in enforcing repayment and reducing the moral hazard problem, also because switching from one e-commerce platform to another could have high costs.[3] By contrast, banks may not be in a position to do likewise as the borrower could have accounts with other banks. Given network effects and high switching costs, big techs could also enforce loan repayments by the simple threat of a downgrade or maybe an exclusion from their ecosystem if in default.

The aim of this paper is to address the following three questions. First, do big tech and bank credit react differently to collateral value, local economic conditions, and firm-specific characteristics? Second, are there differences in the cyclical properties of credit granted to firms that operate in the ecommerce platform (online) and credit granted to firms that operate on traditional business channels (offline)? Third, what are the implications for the monetary transmission mechanism?

To answer these questions, we use a unique dataset that compares the characteristics of loans provided by MYbank, one of the brands under Ant Group, with loans supplied by traditional Chinese banks. In particular, we analyze the credit provided to a random sample of more than 2 million Chinese firms in the period 2017:01–2019:04, extending also our sample to 2020:03 for some structural change tests during the first wave of the Covid-19 pandemic. Differently from the previous literature, the sample of firms used in our study contains not only firms on Alibaba's e-commerce platforms (online firms) but also those that use more traditional business channels (offline firms). The latter use the Alipay app for mobile payments, through the so-called Quick Response (QR) code, but are not fully integrated into the e-commerce platform.

From Ant Group, we obtain access to detailed information on credit supplied by MYbank[4] and firm characteristics on a monthly frequency. In particular, we have access to credit data (quantity and price), and specific information used to model firms' creditworthiness, such as vendor transaction volumes and their network score. The latter measures users' centrality in the network and is based on their payments history and social interactions in the Alipay ecosystem. These pieces of information are then combined with the bank credit history of the client, where we distinguish between secured (backed by collateral) and unsecured (without collateral) bank loans. The comparison between big tech and unsecured bank credit is particularly relevant. To our knowledge, this is the first study that compares the characteristics of big tech credit with bank credit for the same set of firms.

The main results of the paper are the following. First, big tech credit does not correlate with local business conditions and house prices, but reacts strongly to firm characteristics,

3 The switching costs for "vendor/borrowers" who have not repaid their debt could be high for at least three reasons. First, their accounts will be blocked, and they will not be able to use the payment system or their QR codes. Second, e-commerce platforms are quite specialized and moving from one platform to another does not allow the same goods and services to be sold. Third, it takes time to rebuild a reputation.

4 All the data remained located at the Ant Group headquarters and the regression analysis was conducted onsite without the need to export the raw data.

such as transaction volumes and the network score that are used to calculate firm credit ratings. By contrast, both secured and unsecured bank credit react significantly to local house prices, which incorporate useful information on the business conditions in which clients operate and on their creditworthiness. Second, big tech credit to online firms, fully integrated in the e-commerce platform, is more strongly correlated with transaction volumes and network scores than it is in the case of offline firms. Big tech credit to offline firms show some sign of correlation with local demand conditions. Third, an increased use of big tech loans—granted on the basis of big data analysis rather than the use of collateral—makes the supply of credit less reactive to asset price changes and weakens the "collateral channel" (Kiyotaki and Moore, 1997). At the same time, as big tech credit is particularly reactive to firms' specific conditions, a change in their activity will be immediately reflected in the supply of credit, amplifying the effectiveness of the interest rate channel.

The rest of the paper is organized as follows. Section 2 provides a concise literature review and discusses the contribution of our paper. Section 3 presents the data and describes some stylized facts. Section 4 explains our empirical strategy and how we tackle identification issues. Section 5 presents the main results and robustness tests, including an analysis of the stability of the effects during the first wave of the Covid-19 pandemic and a quantification of the real effects of big tech credit provision on firms' business volumes. The last section summarizes the main conclusions.

2. Related Literature

We contribute mostly to three broad strands of literature. First, we provide new supportive evidence on the characteristics of big tech credit, the way it could contribute to increasing financial inclusion and how it could improve risk assessment. Overall, the evidence suggests fintech is growing where the current financial system is not meeting demand for financial services. For the case of China, Hau *et al.* (2021) show that fintech credit mitigates supply frictions (such as a large geographic distance between borrowers and the nearest bank branch) and allows firms with a lower credit score to access credit. In the USA, Tang (2019) finds that fintech credit complements bank lending for small-scale loans. Jagtiani and Lemieux (2018b) find that Lending Club has penetrated areas that are underserved by traditional banks. In Germany, De Roure, Pelizzon, and Tasca (2016) find that fintech credit serves a slice of the consumer credit market neglected by German banks. Cornelli *et al.* (2020) find that fintech and big tech credit are higher where banking sector mark-ups are higher, where there are fewer bank branches, and where banking regulation is less stringent. These papers do not analyze the specific role of data in substituting for collateral in credit provision nor the implications for the monetary transmission mechanism.

Second, we contribute to the empirical literature that studies "agency issues" in the credit market. One mechanism is the so-called "limited commitment" à la Kocherlakota (1996) in environments where risk-sharing arrangements are subject to limited commitments with public and private storage technologies. Ai *et al.* (2021) provide a unified model of dynamic contracting and assortative matching to explain firm dynamics in this context. Another "agency issue" mechanism, intertwined with "limited commitment," can develop in a situation of "missing markets." For example, Lorenzoni (2008) finds that in competitive financial markets, contracts can result in excessive borrowing ex ante and excessive volatility ex post. The inefficiency is due to the combination of limited commitment in financial contracts and the fact that asset prices are determined only in a spot market. Along similar lines, He and Kondor (2016) show that if certain markets are missing, firms' optimal liquidity management could lead to socially inefficient boom-and-bust patterns. In this stream of the literature, collateral plays a key role in mitigating the financial constraints for the development of economic activity (Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997; Bernanke, Gertler and Gilchrist, 1999). Gan (2007) shows that the value and the redeployment ability of collateral affect real-estate prices and corporate investment. Schmalz, Sraer, and Thesmar (2017) find that an increase in collateral value (proxied by house price) leads to a higher probability of becoming an entrepreneur. Our paper investigates a new mechanism that could reduce financial constraints for SMEs: the use of big data and the presence of network effects rather than collateral could provide a different solution to solve agency problems between the lender and the borrower.

Third, our paper also contributes to the empirical literature on how the collateral channel could affect the macroeconomy (Gertler and Gilchrist, 1994; Cloyne *et al.*, 2018; Jeenas, 2018). The use of data instead of collateral for the analysis of creditworthiness could have important implications for the credit channel and the macroeconomy. One example is the link between asset prices, credit, and the business cycle. A rise in collateral values during the expansionary phase of the business cycle fuels a credit boom, while their subsequent fall in a crisis weakens both the demand and supply of credit, leading to a deeper recession. The "collateral channel" is thought to have been one of the main drivers of the Great Depression (Bernanke, 1983) and also an important factor behind the more recent financial crisis (Mian and Sufi, 2011; Bahaj *et al.*, 2019; Ottonello and Winberry, 2020), especially for small businesses (Adelino, Schoar, and Severino, 2015; Doerr, 2021). Indeed, the great financial crisis has shown that the most serious consequences of the drop in the value of collateral were for SMEs that do not have well-diversified funding conditions (OECD, 2019; Lian and Ma, 2021). Using a structural model, Ioannidou, Pavanini, and Peng (2022) shows that a 40% drop in collateral values would lead almost a quarter of loans to become unprofitable, a reduction of average demand by 16% and a drop in banks' expected profits of 25%. Our paper contributes to this stream of the literature by analyzing how big techs' use of big data for credit scoring could attenuate the link between collateral value (house price) and credit supply.

3. Data and Stylized Facts

The empirical analysis in this paper considers Chinese SMEs that obtained credit from MYbank, one of the brands under Ant Group.[5] For these firms, we also observe all loans provided by traditional banks, and distinguish between collateralized credit (secured bank credit) and uncollateralized credit (unsecured bank credit).

The database is constructed at the firm-month level over the period 2017:01–2019:04.[6] The sample includes more than 2 million firms and has been randomly selected from all firms that had transaction records every month and obtained bank credit since January 2017.

Table I presents the summary statistics, divided into three panels: (a) big tech credit; (b) unsecured bank credit; and (c) secured bank credit. For big tech credit, we have more than

5 More information is provided in the Online Annex.

6 In Section 5.7, we extend the database to 2020:03 to perform some structural change tests during the first wave of the Covid-19 pandemic.

Table I. Summary statistics

The sample period is 2017:01–2019:04. Data for credit are not winsorized. Entrepreneurs' characteristics are taken at the date of the issuing of the loan.

(A) Big tech credit	*N*	Mean	St. Dev.	P25	Median	P75
(i) Firms' characteristics						
MYbank credit used (RMB)	6,803,454	20,565	39,592.	2,000	7,115	20,000
Transaction volume monthly (RMB)	6,803,454	25,814	80,477	473	3,031	15,000
Network score	6,803,454	62	22.2	46	58	76
Online	6,803,454	0.312	0.463	0	0	1
(ii) Entrepreneurs' characteristics						
Age	6,803,454	31	7.4	27	31	35
Income level	6,803,454	2.059	0.810	1	2	3
Gender (male = 1, female = 0)	6,803,369	0.682	0.466	0	1	1
(iii) Economic and financial conditions						
House prices (RMB)	6,803,454	18,449	13,817	9,576	12,209	21,515
GDP (billion RMB)	6,803,454	310.494	240.109	124.200	233.910	452.190
Land supply (%)	6,803,454	2.379	1.873	0.923	2.166	3.149
Mortgage rate (%)	6,803,454	5.430	0.377	5.260	5.600	5.720
(B) Unsecured bank credit						
(i) Firms' characteristics						
Unsecured bank credit (RMB)	379,460	128,675	159,266	27,000	70,000	169,000
Transaction volume (RMB)	379,460	24,542	59,025	978	4,967	19,309
Network score	379,460	68	25.9	48	63	83
Online	379,460	0.170	0.376	0	0	0
(ii) Entrepreneurs' characteristics						
Age	379,451	36	6.7	32	36	41
Income level	379,460	2.146	0.829	1	2	3
Gender (male = 1, female = 0)	379,460	0.778	0.416	1	1	1
(iii) Economic and financial conditions						
House price (RMB)	379,460	18,044	12,733	9,832	14,382	21,157
GDP (billion RMB)	379,460	270.562	205.403	121.630	189.270	352.080
Land supply (%)	379,460	2.196	1.637	0.955	2.055	2.879
Mortgage rate (%)	379,460	5.204	0.447	4.690	5.260	5.600
(C) Secured bank credit						
(i) Firms' characteristics						
Secured bank credit (RMB)	91,316	618,577	791,879	150,000	350,000	770,000
Transaction volume (RMB)	91,316	33,066	80,064	1,202	6,480	25,028
Network score	91,316	71	26.5	51	68	88
Online	91,316	0.139	0.346	0	0	0
(ii) Entrepreneurs' characteristics						
Age	91,316	38.561	7	33	38	43
Income level	91,316	2.195	0.822	1	2	3
Gender (male = 1, female = 0)	91,316	0.743	0.437	0	1	1
(iii) Economic and financial conditions						
House price (RMB)	91,316	14,885	9,673	8,782	11,461	16,423
GDP (billion RMB)	91,316	238.142	177.526	114.900	172.960	318.030
Land supply (%)	91,316	2.810	1.934	1.714	2.544	3.413
Mortgage rate (%)	91,316	5.316	0.420	5.010	5.420	5.680

6.8 million firm-month observations. Most of the 2 million MYbank borrowers have access only to big tech credit and do not have a bank relationship. This indicates that big tech credit could be helpful to increase financial inclusion and reduce credit rationing. However, around 47,000 borrowers also have access to secured bank credit and 120,000 to unsecured bank credit, for around 95,000 and 399,000 observations, respectively. Each panel in Table I includes information on: (i) firms' characteristics; (ii) entrepreneurs' characteristics; and (iii) economic and financial conditions where the firm is headquartered. We winsorized all firm and entrepreneur variables at the 1% and 99% level to eliminate outliers.

3.1 Firms' Characteristics

The enterprise data include transaction volumes and credit data. The latter is the actual credit used by the enterprise in a given month. For robustness, we also run some regressions using the overall amount of credit granted by the big tech. Unfortunately, this information is not available for bank credit.

The median credit volume for big tech borrowers is RMB 7,100 (USD 1,065), reflecting the micro nature of MYbank credit and the short maturity of the contract. Big tech credit is typically granted for short periods (from 1 month to 1 year) and then renewed several times, as far as the credit approval remains in place. Often big tech credit assumes the form of a credit line. The median unsecured bank credit is of RMB 70,000 (USD 10,500). The larger size of the loan could reflect a greater length of the contract (from 1 to 3 years). By contrast, the differences in firm size between big tech and bank credit users are not large. The median monthly transaction volume of firms that use big tech credit is RMB 3,000 (USD 450), while that for firms that also use unsecured bank credit is RMB 5,000 (USD 750). Interestingly, the median firm that uses big tech credit is only slightly less connected in the big tech ecosystem (the network score is 58, against an average of 63 for firms that use unsecured bank credit).[7] All in all, firms that borrow only from MYbank are smaller and less connected than the customers of traditional banks, so that the ex-ante potential risk for MYbank is also higher than that of traditional banks, but the differences are not large.

Firms that use secured bank credit are slightly larger; the median transaction volume is RMB 6,500 (USD 975) and with a higher network score (68). The median bank secured credit is RMB 350,000 (USD 52,500). Given the presence of collateral, this also reflects the fact that these loans are typically associated with more important investment decisions by the firm.

3.2 Entrepreneur Information

For SMEs, information about the entrepreneur (typically the owner of the firm or the store) is very important for risk assessment. On the one hand, SMEs have a short life cycle, so the firm information might not be adequately accumulated. On the other hand, the financial situation of SMEs tends to relate very closely to those of the owner.[8] One of the advantages

7 To mitigate concerns about differences in contractual characteristics and firm size, we have run regressions for firms that use all forms of credit, including in the models both time*credit type-fixed effects and borrower*credit type-fixed effects (see Section 5.5). This allows us to control for the possibility that the relationship between a firm and the big tech is different with respect to the relationship between the same borrower and the bank.

8 It is common for SME owners to pledge personal assets to finance their firms (Bahaj, Foulis, and Pinter, 2020). According to the China Micro and Small Enterprise Survey, 92% of secured bank loans to SMEs are collateralized by the entrepreneurs' land or houses.

of MYbank in providing risk control measures for SMEs is that Ant Group can obtain the firm information as well as the information of owners. In this paper, we are able to merge these two different sets of information. Borrowers who access big tech credit are slightly younger (the median age is 31 years) than the owners of firms that use unsecured bank credit (36 years) or secured bank credit (38 years).

Another relevant item of information is the borrowers' level of income. This information is not directly observed by Ant Group, but can be inferred by the total amount of deposits into the Alipay wallet. In particular, we have used this proxy to split the borrowers into three groups: 1 = low income; 2 = medium income; and 3 = high income. The level of income is not very different among the borrowers of the three credit groups: it is on average slightly lower for big tech credit (2.06), intermediate for bank unsecured credit (2.15), and higher for bank secured credit (2.20). More interestingly, the share of female borrowers is significantly higher for big tech credit (32%), than for unsecured and secured bank credit (respectively, 22% and 26%).

The main variables of interest used in this paper are firms' transaction volumes and their network scores, which have a crucial role in the credit scoring analysis of MYbank. These variables are time-varying and can be used in our econometric model, together with borrower-fixed effects.

Figure 1 reports the unconditional elasticity between credit and transaction volume from a random sample of 100,000 firms served by both MYbank and traditional commercial banks. The figure is divided into three panels: the left-hand panel plots big tech credit, the middle panel plots unsecured bank credit, while the right-hand panel plots secured bank credit. Linear trend lines are reported in each graph, together with 95% confidence bands. Interestingly, the elasticity is 0.15 for big tech credit, 0.12 for unsecured bank credit, and 0.09 for secured bank credit, in line with the intuition that big tech credit is more responsive to changes in a borrower's business conditions because banks observe transaction volumes with less precision and with a lag.

Figure 2 evaluates the elasticity between big tech credit and the transaction volume, distinguishing between online and offline borrowers. The yellow dots and the yellow line indicate the offline borrowers (those with a QR code, but not trading in the e-commerce platform), while the blue dots indicate online borrowers (those integrated in the e-commerce platform). The elasticity is 0.090 for offline borrowers and 0.407 for online borrowers. The difference reflects the fact that the big tech firms are able to efficiently collect and process information from online lenders that are integrated in the big tech ecosystem. Therefore, they have access to a rich set of additional data to be combined with traditional transaction volumes obtained from payments. It is interesting to note that the scatterplots can hide two (or more) clouds because Ant Group offers different credit products that target also different clients. For this reason, in Section 5.2, we analyze separately two different products.

Figure 3 evaluates the unconditional elasticity between the three different forms of credit and the network score. The network score is calculated to measure users' centrality in the big tech ecosystem on the base of payment data, users' financial investments, and social interactions.[9] It is worth stressing that both offline and online vendors have a network score

9 The network score is obtained as a rank calculated using a PageRank algorithm. This algorithm was introduced by Larry Page, one of the founders of Google, to evaluate the importance of a particular website page. The calculation is done by means of webgraphs, where webpages are nodes

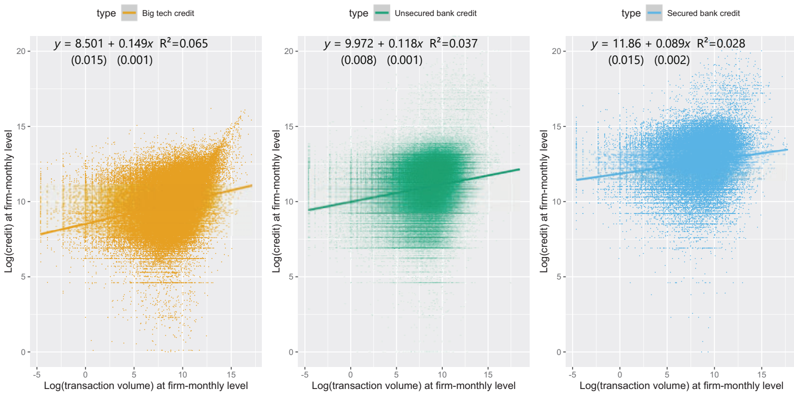

Figure 1. Elasticity between credit and transaction volumes. Based on a 100,000 random sample of firms served by both MYbank and traditional banking. The dots in the figure indicate the log of credit use (y-axis) and the log of transaction volume (x-axis) at the firm-month level. The left-hand panel plots big tech credit, the middle panel plots secured credit, and the right-hand panel plots unsecured bank credit. Linear trend lines are reported in the graphs, together with 95% degree confidence bands. Standard errors in brackets.

because payment data and social interaction information are obtained from Alipay. A user with more connections in the big tech ecosystem has a higher network score. Here, too, the left-hand panel plots big tech credit, the middle panel plots unsecured bank credit, and the right-hand panel plots secured bank credit. The elasticity is 0.83 for big tech credit, 0.30 for unsecured credit, and 0.28 for secured credit. This is not surprising because the network score is not directly observed by the bank and proxies (other) soft information obtained by the bank credit officer on the firm.

Figure 4 plots the correlation between credit and the network score, but distinguishing between offline and online firms. The yellow dots and the yellow line indicate the offline borrowers (those with a QR code, but not integrated in the e-commerce platform), while the blue dots indicate the online borrowers (those perfectly integrated in the ecommerce platform). Credit reveals a positive and significant elasticity with network effect that is approximately same for online and offline borrowers (respectively, 1.120*** and 1.187***). This preliminary evidence shows that the network measure is extremely important for credit

and hyperlinks are edges. Each hyperlink to a page counts as a vote of support for that webpage. In the case of the Ant Group network score, customers and QRcode merchants can be considered as interconnected nodes (webpages) and payment funding flows can be considered as edges (hyperlinks). There are several reasons why more connections in the big tech ecosystem translates into higher creditworthiness. (1) Anti-fraud: the transaction records of the user with a high network score means that the user is more reliable. The predictive power of the network score system consists in exploiting the network structure between vendors and customers. For instance, fraudulent applications (fake businesses) are detected by identifying isolated clusters of nodes that have limited connections with other businesses. (2) Social network: the connections can reflect the popularity of the user and his products. (3) Big tech ecosystem value: the big tech ecosystem is more valuable (more ecommerce capacity) for users with a higher network score.

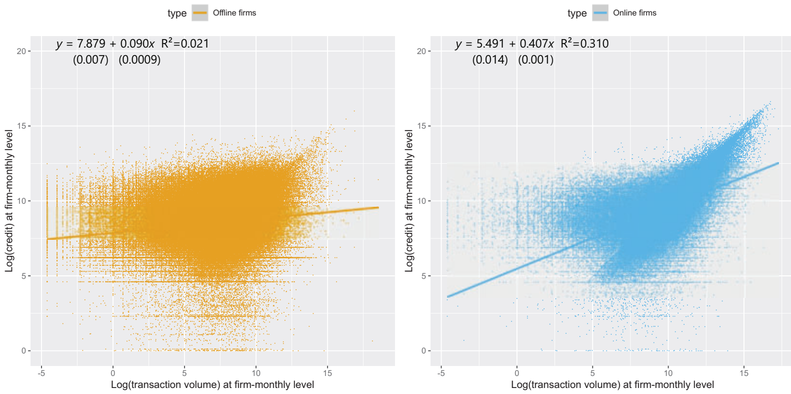

Figure 2. Elasticity between big tech credit and transaction volume: offline firms versus online firms. Based on a 100,000 random sample of firms that received credit by MYbank. The dots in the figure indicate the log of credit use (*y*-axis) and the log of transaction volume (*x*-axis) at the firm-month level. The left-hand panel plots credit to offline firms and the right-hand panel plots credit to online firms. Linear trend lines are reported in both graphs, together with 95% degree confidence bands. Standard errors in brackets.

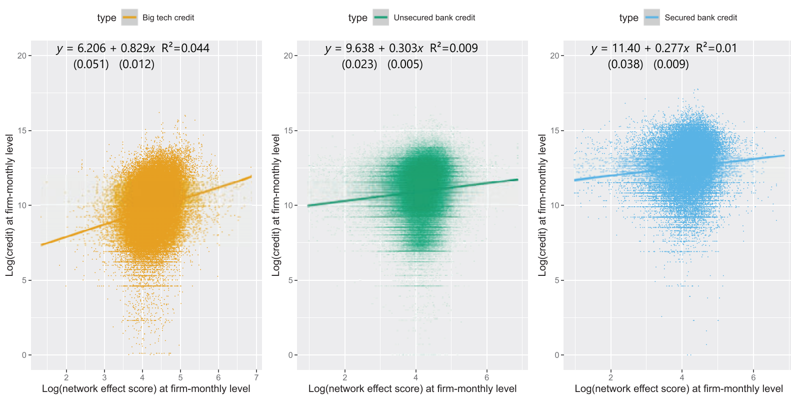

Figure 3. Elasticity between credit and the network score. Based on a 100,000 random sample of firms served by both MYbank and traditional banks. The dots in the figure indicate the log of credit use (*y*-axis) and network score (*x*-axis) at the firm-month level. The left-hand panel plots big tech credit, the middle panel plots secured credit, and the right-hand panel plots unsecured bank credit. Linear trend lines are reported in the graphs, together with 95% degree confidence bands. Standard errors in brackets.

scoring, also for those firms that do not conduct their main activity in the e-commerce platform.

3.3 House Prices, GDP, and Monetary Policy

The data source on house prices is the 100-city housing prices published by China Index Academy and included in the WIND database. The data cover 100 samples of new houses

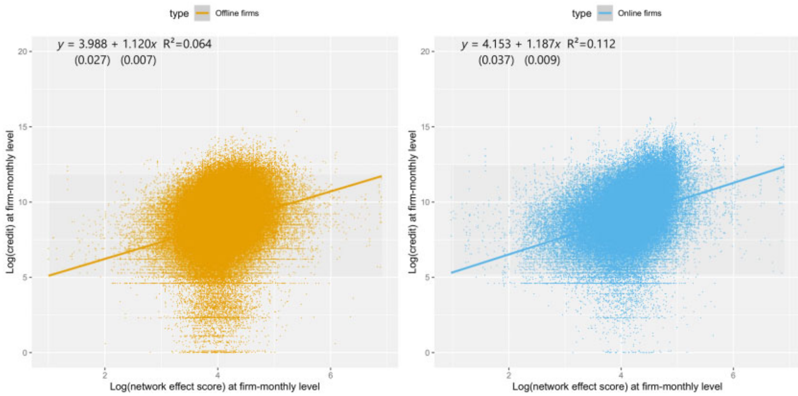

Figure 4. Elasticity between big tech credit and the network score: offline firms versus online firms. Based on a 100,000 random sample of firms that received credit by MYbank. The dots in the figure indicate the log of credit use (y-axis) and the log of network effect score (x-axis) at the firm-month level. The left-hand panel plots credit to offline firms and the right-hand panel plots credit to online firms. Linear trend lines are reported in both graphs, together with 95% degree confidence bands. Standard errors in brackets.

for sale in China, including commercial housing, villas, and affordable housing, and all the houses for sale with a sales license were included in the calculation.[10]

Figure 5 indicates the unconditional elasticity between the different credit forms and house price. The dots in the figures indicate the average logarithm credit use (y-axis) and the average logarithm of housing price (x-axis) at the city-year level. The left-hand panel plots big tech credit, the middle panel plots bank unsecured credit, and the right-hand panel plots bank secured credit. Linear trend lines are reported in each graph, together with 95% confidence bands. The (unconditional) elasticity of big tech credit with respect to house prices is 0.09, while that of unsecured bank credit is twice as high (0.184) and that of secured bank credit is five times higher (0.488).

The different elasticities of the three credit types with respect to house prices remain quite stable even controlling for different local GDP conditions and including a complete set of time-fixed effects. Figure 6 reports the different elasticities and associated standard errors for this simple model. Interestingly, the elasticity of unsecured bank credit with respect to GDP at the city level is more than four times higher (0.032) than that of big tech credit (0.007). The elasticity of bank secured credit with respect to local GDP is not statistically different from zero.

3.4 Credit Quality and Interest Rates

Big tech credit has lower default rates than bank credit. Table II compares non-performing loans (NPLs) for Chinese banks and for MYbank, focusing on credit to SMEs. As reported in the first two rows of the table, NPLs for the Chinese banking industry have been substantially higher on average than for MYbank in the period under investigation in this paper

10 According to the available data, the 100-city housing price data are the database with the largest coverage of monthly housing prices in China. China Index Academy has published the *China Real Estate Statistical Yearbook* for 16 consecutive years with the State Statistics Bureau.

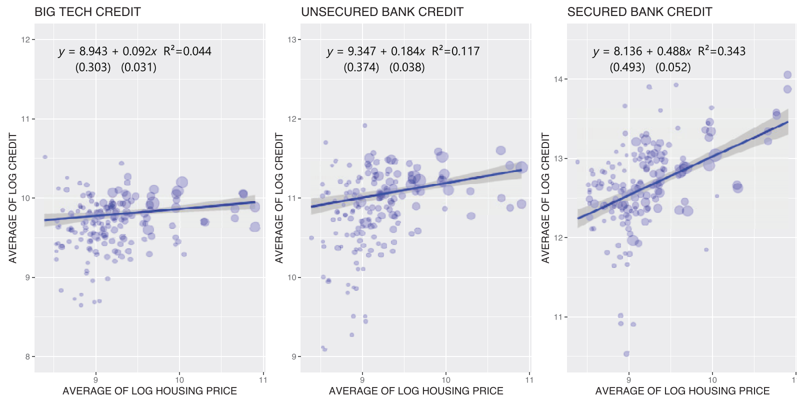

Figure 5. Elasticity of credit with respect to house prices. Based on a 100,000 random sample of firms served by both MYbank and traditional banking. The dots in the figures indicate the average logarithm credit use (*y*-axis) and the average logarithm of housing price (*x*-axis) at the city-year level. Growth rates are approximated using first differences of log values. The left-hand panel plots big tech credit, the middle panel plots bank secured credit, and the right-hand panel plots bank unsecured credit. Linear trend lines are reported in each graph, together with 95% degree confidence bands. Standard errors in brackets.

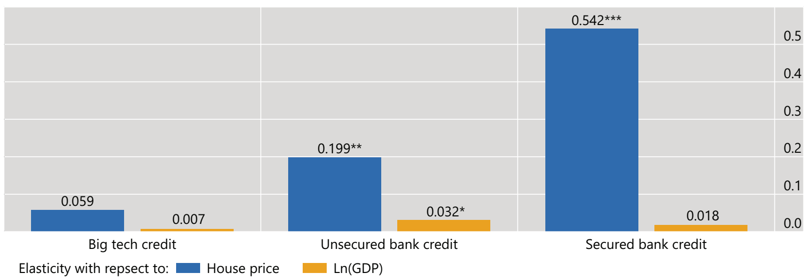

Figure 6. Elasticity of credit with respect to house prices and GDP. The figure reports the coefficient of three different regressions (one for each credit types) in which the log of credit is regressed with respect to the log of house prices at the city level, the log of GDP at the city level, and a complete set of time dummies. Significance level: $^{**}P < 0.05$; $^{***}P < 0.01$.

(2017–9) and also during the Covid-19 pandemic (2020). These results are consistent with Huang *et al.* (2020), who find that big tech credit scoring yields better prediction of loan defaults during normal times and periods of large exogenous shocks, reflecting information and modeling advantages.

Interestingly, the ex-post measure of credit risk is not mirrored in the interest rates that are substantially higher (on average) for big tech credit. Three reasons may cause interest rates for big tech credit to be higher than those for bank credit. First, the funding costs of MYbank are substantially higher than those of traditional banks. This reflects big techs' limited ability to accept retail deposits. Big techs could potentially establish an online bank, but regulatory authorities typically restrict the opening of remote (online) bank accounts. One relevant example is China, where the two Chinese big tech banks (MYbank and

Table II. Credit quality and interest rates

NPLs indicate loans that are typically overdue from 90 days and more. See "Interim Measures for the Risk Classification of Financial Assets of Commercial Banks 商业银行金融资产风险分类暂行办法." (1) Credit lines below 10 million Yuan (5 million in 2017 and 2018). (2) Data obtained from public balance sheet information dividing interest earned and total loans for SMEs. (3) January–August 2020. (4) January–May 2020.

Source: CBIRC, Annual reports of MYbank.

	Credit quality SMEs: NPL ratio		Average interest rates SMEs	
Years	Banks (1)	MYbank	Banks (1)	MYbank (2)
2017	5.85%	1.23%	6.55%	17.70%
2018	5.50%	1.30%	6.16%	13.39%
2019	3.22%	1.30%	6.70%	10.21%
2020	2.99% (3)	1.52%	6.03% (4)	9.03%

WeBank) rely mostly on interbank market funding and certificates of deposit that are typically more costly than retail deposits (BIS, 2019). Second, as discussed above, firms that borrow from MYbank are smaller than the customers of traditional banks, so the ex-ante potential risk for MYbank is also higher than that of traditional banks. Third, data processing for credit scoring could have high fixed costs to set up the necessary IT infrastructure and create a highly specialized team. These costs could be particularly high at the beginning, when the number of borrowers is low, and then decline with time, when the market share increases. Interestingly, this is reflected by the spread between big tech credit and bank credit interest rates that were around 10% in 2017, when MYbank started to offer credit to QR code merchants, and only 3% in 2020.

4. Econometric Strategy

Our analysis starts with a simple model that analyses the main determinants of credit. We consider the following baseline model:

$$\ln\left(\text{credit}_{i,j,t}\right) = \mathrm{A}^{'} X_{i,j,t} + \Gamma^{'} Y_{j,t} + \mu_i + \mu_t + \varepsilon_{i,t}, \tag{1}$$

where $\ln\left(\text{credit}_{i,j,t}\right)$ is the logarithm of the credit granted by MYbank or traditional banks (unsecured and secured) to firm i, headquartered in city j, in time t. $X_{i,j,t}$ is a vector that contains time-variant firm characteristics (transaction volume and network score). $Y_{j,t}$ are the city-level indicators to capture regional conditions, including log of house price and local GDP.[11] This model includes firm- (μ_i) and time-fixed effects (μ_t), while $\varepsilon_{i,t}$ is an error term. Following Chaney, Sraer, and Thesmar (2012), we cluster the standard errors at the city-month level.

This model is able to control for all firm (unobserved) invariant characteristics but at the cost to limit the analysis to firms that received at least two loans in the period under

11 Giving the quarterly frequency of local GDP, different from the monthly frequency of all the other variables, we have lagged the log of GDP of one period. This choice should mitigate endogeneity concerns.

investigation. Due to the inclusion of firm-fixed effects, we cannot include in vector $X_{i,j,t}$ the time-invariant entrepreneur's characteristics (age, gender, and income).[12]

We run this model for the three different types of credit and compare the coefficients of housing price and local GDP to evaluate if big tech and bank credit (unsecured and secured) react differently to business and asset price conditions. Moreover, we can distinguish the effects between offline firms and online firms in order to verify the additional effects—if any—for firms that are fully integrated into the big tech ecosystem. As big tech credit could take different contractual forms, we also run model (1) within homogenous credit products categories to check for the possible existence of aggregation biases.

Another identification challenge derives from the fact that the customers of the big tech company could be very different from those of a traditional bank and it could be in principle difficult to compare them. To address this concern, we select the firms that have both big tech credit and traditional bank credit and use a difference-in-difference approach that follows Khwaja and Mian (2008). This approach allows us to compare the characteristics of credit from different sources used by the same customer. In particular, we used a nested model in which big tech credit and bank credit (unsecured or secured) are jointly analyzed. In this case, we include both time * credit type (μ_{tc})-fixed effects and borrower * credit type (μ_{ic})-fixed effects. The inclusion of these fixed effects controls for the fact that the relationship between a borrower and the big tech firm could be quite different with respect to the relationship between the same borrower and the bank. The inclusion of these fixed effects necessitates for each firm to have at least two big tech credits and two bank credits over the sample horizon. In particular, we run the following model:

$$\ln\left(\text{credit}_{i,j,t}\right) = \mathrm{A}X_{i,j,t} + \Gamma Y_{j,t} + BX_{i,j,t} * \text{credit\_type} + KY_{j,t} * \text{credit\_type} + \mu_{ic} + \mu_{tc} + \varepsilon_{i,t}. \tag{2}$$

The different reaction of bank credit with respect to big tech credit is evaluated by interacting a credit_type dummy that takes the value of 1 for bank unsecured (or bank secured) credit and 0 for big tech credit. The test for the difference between the coefficients is given directly by the sign and the significance of the interaction terms (B' for the borrower-specific characteristics and K' for the local economic conditions).

Another identification challenge is the necessity to properly control for demand shifts. Model (2) can be further enriched to control for specific changes in the economic conditions at the city level that could affect the credit market. One possibility is to integrate the model with city * time-fixed effects (μ_{jt}) to control for changes in local conditions over time. In doing so, however, the local economic indicators are subdued $Y_{j,t}$ by the city * time-fixed effect. The model becomes

$$\ln\left(\text{credit}_{i,j,t}\right) = \mathrm{A}X_{i,j,t} + BX_{i,j,t} * \text{credit\_type} + KY_{j,t} * \text{credit\_type} + \mu_{ic} + \mu_{tc} + \mu_{jt} + \varepsilon_{i,t}. \tag{3}$$

In this model, we can focus our attention on the significance of the interaction terms B' (or K') to evaluate a different reaction of credit types to borrower-specific characteristics (local economic conditions).

12 We also analyze a model that includes time-invariant firm characteristics and exclude firm-fixed effects. The results (not reported for the sake of brevity but available upon request) are very similar.

Following Jimenez *et al.* (2014), an alternative way to control for shifts in demand is to progressively saturate Model (2) with Time * Borrowers (μ_{ti})-fixed effects, together with City*credit_type (μ_{jc}) and Time*credit_type (μ_{tc})-fixed effects. In this way, we use borrower*time to absorb all time-varying, observed and unobserved firm heterogeneity. City*credit_type and Time*credit_type control the location-varying and time-varying heterogeneity of different credit types. This specification allows us to control more precisely for borrower-specific demand shifts but necessitates a further restriction of the sample to consider only firms that have in place both big tech credit and bank credit in one month. We estimate the following regression:

$$\ln\left(\text{credit}_{i,j,t}\right) = \mathrm{A}X_{i,j,t} + \Gamma Y_{j,t} + BX_{i,j,t}{*}\text{credit\_type} + K\mathrm{Y}_{j,t}{*}\text{credit\_type} + \mu_{ti} + \mu_{jc} + \mu_{tc} + \varepsilon_{i,t}. \tag{4}$$

A final concern is reverse causality. In principle, credit expansion by (large) firms may also affect house prices; however, we argue that this concern is unlikely to affect our results because the firms in our sample are relatively small. Following Chaney, Sraer, and Thesmar (2012), we also instrument the house price by using a hand-collected monthly measure for land supply by the Government and its interaction with mortgage rates in order to insulate price movements that are (exogenously) driven by supply changes.

5. Results

5.1 Baseline Model

Table III presents the results of Model (1). The left-hand panel of the table reports the results when the log MYbank credit used by the borrowers is considered as dependent variable. Big tech credit is not correlated with house prices and local economic conditions. By contrast, MYbank credit used by the borrowers is strongly correlated with firm-specific variables (transaction volumes and network score), reflecting the nature of the financial services provided by big techs that is tailored toward the characteristics of their clients. The correlation is lower for firms that work offline than for firms that work online. The latter are indeed more integrated in the big tech ecosystem and MYbank could get more information on them. This result indicates a greater capacity for data obtained in the e-commerce platform to capture (otherwise) unobservable firms' characteristics and alleviate adverse selection problems. The fit of these three regressions is quite good: the adjusted R^2 indicates that the models for the three different sample splits are always able to capture around 60% MYbank credit variability.

The analysis presented so far has considered as a dependent variable the credit used by the firm. In order to grasp more information on the credit lines supplied by MYbank we consider in the right-hand panel of Table III, the value of the credit line granted by MYbank as a dependent variable. Credit lines represent around 20% of total credit contracts and indeed the number of observations drops from 6.8 to 1.3 million.

The results remain similar: local GDP and house prices are not significantly correlated with big tech credit lines granted, both for offline and online borrowers. As now the dependent variable represents big tech credit supply, the result reinforces the interpretation that big tech credit assessment focuses more on firm-specific information rather than on local economic conditions (which more closely reflect changes in the demand for credit).

Table III. Drivers of big tech credit

The sample period is 2017:01–2019:04. (1) At the city-month level. (2) At the city-quarter level. Lagged one period. (3) Network score measures users' centrality in the network and is based on users' payment and funds information and social interactions. The user who has more connections gets a higher network score. Standard errors reported in brackets are clustered at the city-month level. Significance level: *$P<0.1$; **$P<0.05$; ***$P<0.01$.

Explanatory variables	Dependent variable: Log (MYbank credit used)			Dependent variable: Log (MYbank credit line granted)		
	All (I)	Offline (II)	Online (III)	All (IV)	Offline (V)	Online (VI)
Log house price (1)	0.005	0.060	0.012	−0.027	−0.065	0.137
	(0.039)	(0.049)	(0.042)	(0.066)	(0.068)	(0.102)
Log GDP (2)	0.007	0.008	−0.006	−0.002	−0.003	0.00004
	(0.006)	(0.006)	(0.007)	(0.003)	(0.004)	(0.006)
Log transaction volume	0.046***	0.021***	0.159***	0.0005**	0.0004*	0.001*
	(0.001)	(0.0004)	(0.002)	(0.0002)	(0.0002)	(0.001)
Log network score (3)	0.578***	0.091***	0.863***	0.219***	0.209***	0.189***
	(0.013)	(0.010)	(0.011)	(0.010)	(0.011)	(0.026)
Time FE (month)	Yes	Yes	Yes	Yes	Yes	Yes
Borrower FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	6,803,454	4,689,936	2,120,518	1,256,305	1,104,803	151,502
Adjusted R-squared	0.616	0.628	0.590	0.918	0.916	0.930

5.2 Homogenous Big Tech Credit Contracts and Interest Rates

The results presented in the first three columns of Table III refer to the overall value of big tech credit used, aggregating credit contractual forms that could be quite different. It is important therefore to replicate the analysis considering specific and more homogenous forms of credit contracts. This will allow us to test for the possible existence of biases related with the aggregation of credit with different contractual conditions.

In Table IV, therefore, we report the regressions for Model (1) for two particular products offered by MYbank to firms. The left-hand side of the table considers a very popular credit product (Product 1) that is directly accessible to the firm in very simple steps, using a smartphone for instance. The application for the credit is completed with a few taps on the screen and no collateral is required. This contractual form is particularly used by offline firms (QR code merchants). MYbank offers a credit line for each merchant that is based on her specific risk profile. As long as the (offline) merchants use Alipay QR code to collect payments, they will have the opportunity to obtain and renew the credit.

The second credit contract (Product 2) is offered by MYbank to firms on the basis of the overall value of their orders and receivables in the Taobao platform. As the information is obtained on the e-commerce platform, this contractual form is used by online firms. This product is a trade credit product. Every vendor can access a credit line from MYbank, but the amount of credit granted to each customer is determined by accounts receivable.

Table IV is divided into two parts. The first two columns consider as dependent variables the quantity of credit, while the last two columns report results for regressions where the interest rate is the dependent variable. Neither form of big tech credit is correlated with

Table IV. Drivers of big tech credit: two specific products

The sample period is 2017:01–2019:04. (1) At the city-month level. (2) At the city-quarter level. Lagged one period. (3) Network score measures users' centrality in the network and is based on users' payment and funds information and social interactions. The user who has more connections gets a higher network score. Standard errors reported in brackets are clustered at the city-month level. Significance level: $^{*}P<0.1$; $^{**}P<0.05$; $^{***}P<0.01$.

Explanatory variables	Dependent variable: Log (credit used)		Dependent variable: Interest rate	
	Product 1	Product 2	Product 1	Product 2
Log house price (1)	0.038	–0.007	–0.0002	0.0003
	(0.111)	(0.116)	(0.001)	(0.002)
Log GDP (2)	0.007	0.003	0.00001	–0.0004
	(0.014)	(0.018)	(0.0001)	(0.0003)
Log transaction volume	0.011***	0.598***	–0.00001***	–0.001***
	(0.001)	(0.004)	(0.00001)	(0.00004)
Log network score (3)	0.091***	0.390***	–0.001**	–0.005***
	(0.035)	(0.024)	(0.0003)	(0.0003)
Time FE (month)	Yes	Yes	Yes	Yes
Borrower FE	Yes	Yes	Yes	Yes
Number of observations	705,542	179,500	714,832	181,516
Adjusted R-squared	0.364	0.690	0.916	0.793

house prices nor with local economic conditions. By contrast, they are highly correlated with borrower-specific characteristics (both transaction volumes and network score). The correlation is higher for Product 2, used by firms that work online, while the correlation for Product 1, used by firms that work offline, is lower.

When considering the interest rate as the dependent variable (see the third and fourth columns of Table IV), the results parallel those obtained on quantities. Interest rates do not react to the evolution of house prices and local economic conditions. By contrast, price conditions react to borrowers' specific characteristics, more strongly for Product 2 that is available for firms that operate online on the big tech e-commerce platform.

5.3 Main Drivers of Bank Lending: Collateralized versus Uncollateralized Contracts

Table V reports the result of Model (2) for bank credit. Unsecured bank credit is correlated with house prices but the elasticity is significantly lower than for secured bank credit (the elasticities are 0.203 and 0.543, respectively). The positive correlation between unsecured bank credit and house prices could reflect higher demand in cities with higher asset prices, with the latter reflecting in general better economic conditions. We will try to filter out this effect in Section 5.6. Another explanation may be that banks do not have enough granular information on the firm, so local house price dynamics turn out to be one relevant indicator

Table V. Drivers of bank credit

The sample period is 2017:01–2019:04. (1) At the city-month level. (2) At the city-quarter level. Lagged one period. (3) Network score measures users' centrality in the network and is based on users' payment and funds information and social interactions. The user who has more connections gets a higher network score. Standard errors reported in brackets are clustered at the city-month level. Significance level: *$P<0.1$; **$P<0.05$; ***$P<0.01$.

Explanatory variables	Dependent variable: Log (unsecured bank credit)			Dependent variable: Log (secured bank credit)		
	All	Offline	Online	All	Offline	Online
Log house price (1)	0.203***	0.242***	0.014	0.543***	0.439***	1.037***
	(0.085)	(0.088)	(0.178)	(0.149)	(0.152)	(0.337)
Log GDP (2)	0.030*	0.029*	0.041	–0.032	–0.034	–0.027
	(0.017)	(0.017)	(0.034)	(0.033)	(0.032)	(0.105)
Log transaction volume	0.004***	0.003***	0.005*	0.003	0.001	0.014**
	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)	(0.007)
Log network score (3)	0.023	0.028*	–0.017	–0.041	–0.048	–0.046
	(0.015)	(0.017)	(0.040)	(0.030)	(0.033)	(0.086)
Time FE (month)	Yes	Yes	Yes	Yes	Yes	Yes
Borrower FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	379,460	314,817	64,643	91,316	78,601	12,715
Adjusted R-squared	0.639	0.638	0.640	0.567	0.568	0.562

to identify a firm's creditworthiness. Interestingly, the correlation is not statistically significant for firms that operate online and for which local conditions are less relevant. Unsecured bank credit is also positively correlated with local GDP conditions for reasons similar to those discussed above for house prices.

Unsecured bank credit is also correlated with borrower-specific characteristics, especially for offline borrowers. This could reflect the fact that online vendors' activity is less visible to banks than that of offline vendors (e.g., a restaurant or a shop). The physical presence in the territory could indeed be relevant for a bank credit officer who could observe more directly firms' characteristics. Collateralized bank credit shows some signs of correlation only with respect to transaction volumes for firms that work online. However, we will see later that this result vanishes when more complete specifications are used that control for demand shifts.

5.4 Endogeneity Issues

In principle, there are potential sources of endogeneity in Model (1) and house prices could affect credit through channels other than rising collateral values. This could happen for three reasons. The first one is a simple reverse causality argument: large firms may have a non-negligible impact through the demand for local labour and locally produced goods on local activity, so that an increase in credit demand for such large firms could trigger also a housing price appreciation. This would lead us to overestimate the coefficient on housing price. Second, it could be that our measure of housing prices proxies for local demand shocks that are not fully captured by local GDP conditions. Third, expansion in credit may also have effects on house prices (Favara and Imbs, 2015).

The first issue is unlikely to affect our results because the firms analyzed in this study are of small dimension and their credit decisions are unlikely to affect local output via increase in local labor and/or increase in produced goods. Furthermore, we have winsorized all the firms and entrepreneurs' variables at the 1% and 99% level to eliminate the effects of outliers. On the second issue, we have used city*time-fixed effect and borrower*time-fixed effect to control for shifts in the demand side. The inclusion of borrower*credit_type in nested models will allow us to control for a heterogeneous demand schedules for big tech and bank credit for the same client (see Section 5.5).

To address the third issue, we instrument the housing price. Panel A in Table VI presents the results of the first-stage regression where we use one-year lagged land supply and its interaction with mortgage rates as instrumental variables. The local government has a great influence on housing prices through the land supply in China (Glaeser *et al.*, 2017). The literature on the determinants of house prices for China indeed uses information about land as an instrumental variable to model housing price. For example, Hau and Ouyang (2018) use the lagged value of the surface of newly useable residential land scaled by the size of the existing housing stock and local population density. Waxman *et al.* (2020) use the lagged volume (in square kilometers) of cumulative land sales in each city.

In line with these papers, we use a hand-collected monthly measure for land supply. In particular, we have calculated for each month the annual cumulative measure for land supply for each local government scaled by urban construction land. Our measure represents an improvement over the other measures indicated above, such as land sales or the proxies for local government land supply. In our empirical model, following Chaney, Sraer, and Thesmar (2012), we also include the interaction between the land supply measure and mortgage rate as an instrumental variable. This should control for differential price effects in different cities caused by a different sensitivity to monetary policy conditions. The results indicate that, as expected, land supply has a negative effect on house prices. When mortgage rates decrease, house prices of cities with higher land supply increase by less. There may be one concern about the endogeneity of the mortgage rate. In particular, mortgage rates could be correlated to local conditions. However, the mortgage rate used in our first-stage regression is nationwide and highly correlated to the benchmark interest rate controlled by the People's Bank of China and hence, for practical purposes, exogenous to local conditions.

Panel B in Table VI presents the results of Model (1) on big tech credit, unsecured bank credit, and secured bank credit, using the log house prices instrumented in Panel A. Only bank credit is significantly correlated with house prices: the elasticity of unsecured bank credit with respect to house prices is 0.408, while that of secured bank credit is 0.690. This result underscores that, in the case of an (exogenous) decrease in the value of collateral triggered by an expansion in the supply of land by the government, there is no positive effect on big tech credit.[13] Other things being equal, these results indicate a reduction in the effects of

13 The elasticity of the three credit types with respect to house price tends to increase with IV regressions, pointing to the fact that the baseline regression estimates could be biased downward. However, qualitatively the results are very similar. The elasticity of big tech credit with respect to house price remains insignificant and the elasticity of secured bank credit increases from 0.54*** to 0.69**. The only noticeable change is for unsecured bank credit whose elasticity with respect to house price doubles (from 0.203*** to 0.408**). As the effects of house prices on unsecured bank credit likely reflect general information on the business conditions in which firms operate and on their creditworthiness, IV estimates seem to filter out better such effect. Indeed,

Table VI. Instrumental variable regressions

The sample period is 2017:01–2019:04. (1) Lagged land supply are calculated using annual land supply scaled by urban construction land lagged by 12 months. (2) Nationwide interest rate at which banks refinance their home loans at the quarterly level. Standard errors in brackets are clustered at the city level. (3) Log house prices are instrumented using the model described in panel (A). (4) At the city-quarterly level. Lagged one period. (5) Network score measures users' centrality in the network and is based on users' payment and funds information and social interactions. The user who has more connections gets a higher network score. Standard errors in brackets are clustered at the city-month level. Significance level: $^{*}P<0.1$; $^{**}P<0.05$; $^{***}P<0.01$.

Panel A. First-stage regression

Explanatory variables	Dependent variable: Log house price
Lagged land supply (1)	–0.0370***
	(0.0131)
Lagged land supply (1) * mortgage rate (2)	0.00667**
	(0.00269)
Time FE	Yes
City FE	Yes
Number of observations	2,688
Adjusted R-squared	0.9952

Panel B. Second-stage regression

Explanatory variables	Dependent variable: Log (MYbank credit used)	Dependent variable: Log (unsecured bank credit)	Dependent variable: Log (secured bank credit)
Log house price IV (3)	0.039	0.408**	0.690**
	(0.081)	(0.199)	(0.330)
Log GDP (4)	0.007	0.028*	–0.031
	(0.006)	(0.017)	(0.033)
Log transaction volume	0.046***	0.004***	0.003
	(0.001)	(0.001)	(0.002)
Log network score (5)	0.577***	0.022	–0.043
	(0.013)	(0.015)	(0.030)
Time FE	Yes	Yes	Yes
Borrower FE	Yes	Yes	Yes
Number of observations	6,803,454	379,460	91,316
Adjusted R-squared	0.616	0.639	0.567

the collateral channel of monetary policy. By contrast, big tech credit is not correlated with local economic but is strongly correlated with firms' transaction volume and network score and this indicates a larger reactivity in this form of credit to changes in firm-specific conditions.

the IV regression controls in the first stage for the general impact of local housing supply and interest rate conditions on housing price.

5.5 Nested Models

As discussed in Section 4, the comparison between the coefficients of Model (1) across the different credit types is difficult because the estimations are derived from groups of firms with different characteristics.

Table VII presents the results of the nested Model (2) in which big tech credit and unsecured bank credit are jointly analyzed. To check for unobservable characteristics, we include in this case both Time*credit_type-fixed effects and borrower*credit_type-fixed effects. The different reaction of each form of credit with respect to the explanatory variables is evaluated by interacting the latter with a dummy variable ("Bank unsecured") that takes the value of 1 for bank unsecured credit and 0 for big tech credit. The test for the difference in the coefficients between the two different forms of credit is given directly by the sign and the significance of the interaction term. For example, fintech credit does not correlate with local GDP conditions (the coefficient is 0.025 with a standard error of 0.018), while unsecured bank credit does ($0.025 + 0.055 = 0.08$), with the difference between fintech credit and unsecured bank credit that is statistically significant at the 10% level (0.055*). Differences in the elasticity of the two forms of credit with respect to house prices are statistically significant when considering offline firms. The other results are qualitatively similar to those already reported.

Table VIII presents the results of the nested Model (3) in which big tech credit and secured bank credit are jointly analyzed. In this case, as well, the different reaction of each form of credit with respect to the explanatory variables is evaluated by interacting the latter with a dummy variable ("Bank secured") that takes the value of 1 for bank secured credit and 0 for big tech credit. The test for the difference in the coefficients between the two different forms of credit is given directly by the sign and the significance of the interaction term. For example, big tech credit does not correlate with house price (the coefficient is -0.002 with a standard error of 0.156), while secured bank credit does ($-0.002 + 0.448 = 0.445$), with the difference between fintech credit and secured bank credit that is statistically significant (0.448***).

The other results are confirmed. Big tech credit and bank secured credit are not correlated with local economic conditions. Big tech credit is highly correlated with borrower-specific characteristics (transaction volumes and network score), more for e-commerce firms that work online. By contrast, bank secured credit is not correlated with borrower-specific characteristics (transaction volumes and network score).

5.6 Additional Controls for Changes in Local Conditions and Demand Shifts

Table IX presents the comparison between big tech credit and bank credit including additional controls for local condition or borrowers' demand shifts. Indeed, one concern for our results is that the evolution of quarterly GDP at the city level is not sufficient to fully capture the effects on firms' demand. This could be particularly important for offline firms that are more affected by local economic conditions. We report therefore in the first two columns of Table IX the results using Equation (3) that includes Time*city-fixed effect (together with Borrower*credit_type and Time*credit_type-fixed effects) to control for unobserved (to the econometrician) change in local conditions.

The last two columns of Table IX consider instead Equation (4) with a complete set of Time*borrower-fixed effects (and also City*credit_type and Time*credit_type-fixed effects). These controls are more stringent, and their inclusion does not allow us to keep in the specification time-varying macroeconomic and borrower characteristics. In this case, we simply

Table VII. Big tech credit versus bank unsecured credit

The sample period is 2017:01–2019:04. (1) At the city-month level. (2) At the city-quarter level. Lagged one period. (3) Network score measures users' centrality in the network and is based on users' payment and funds information and social interactions. The user who has more connections gets a higher network score. (4) Dummy variable that takes the value of 1 for bank unsecured credit and 0 for big tech credit. Standard errors in brackets are clustered at the city-month level. Significance level: *$P < 0.1$; **$P < 0.05$; ***$P < 0.01$.

Explanatory variables	Dependent variable: Log (credit)		
	All	Offline	Online
Log house price (1)	0.001	–0.078	0.120
	(0.074)	(0.085)	(0.122)
Log GDP (2)	0.025	0.026	0.024
	(0.018)	(0.020)	(0.035)
Log transaction volume	0.037***	0.016***	0.114***
	(0.001)	(0.001)	(0.003)
Log network score (3)	0.383***	0.054**	0.745***
	(0.022)	(0.025)	(0.034)
Log house price * bank unsecured (4)	0.096	0.224*	–0.176
	(0.117)	(0.120)	(0.224)
Log GDP * bank unsecured (4)	0.055*	0.057*	0.044
	(0.030)	(0.031)	(0.058)
Log transaction volume * bank unsecured (4)	–0.035***	–0.013***	–0.112***
	(0.002)	(0.002)	(0.004)
Log network score * bank unsecured (4)	–0.361***	–0.024	–0.752***
	(0.028)	(0.032)	(0.051)
Time * credit type FE	Yes	Yes	Yes
Borrower * credit type FE	Yes	Yes	Yes
Number of observations	646,203	481,424	164,779
Adjusted R-squared	0.679	0.691	0.654

focus on the interaction terms between each variable and the credit type dummy. Moreover, using this specification, we need to further restrict the number of observations as the analysis can only be carried out for borrowers who have both big tech credit and bank credit in one month.[14]

Even after controlling more appropriately for demand shifts, (unsecured and secured) bank credit is more correlated than big tech credit with respect to house prices, and the difference is particularly high for the more restrictive model (4) that includes Time*borrower-fixed effects. Controlling for demand shifts, unsecured and secured bank credit is always more correlated with local economic condition than big tech credit. By contrast, the latter

14 It is worth remembering that in the first and third columns, the dummy variable Bank credit takes the value of 1 for bank unsecured credit and 0 for big tech credit. Vice versa, in the second and fourth column, the dummy variable Bank credit takes the value of 1 for bank secured credit and 0 for big tech credit.

Table VIII. Big tech credit versus bank secured credit

The sample period is 2017:01–2019:04. (1) At the city-month level. (2) At the city-quarter level. Lagged one period. (3) Network score measures users' centrality in the network and is based on users' payment and funds information and social interactions. The user who has more connections gets a higher network score. (4) Dummy variable that takes the value of 1 for bank secured credit and 0 for big tech credit. Standard errors in brackets are clustered at the city-month level. Significance level: *$P<0.1$; **$P<0.05$; ***$P<0.01$.

Explanatory variables	Dependent variable: Log (credit)		
	All	Offline	Online
Log house price (1)	–0.002	–0.286	0.249
	(0.156)	(0.196)	(0.238)
Log GDP (2)	–0.010	–0.016	0.002
	(0.025)	(0.028)	(0.050)
Log transaction volume	0.040***	0.015***	0.157***
	(0.003)	(0.003)	(0.007)
Log network score (3)	0.465***	0.106**	0.883***
	(0.045)	(0.048)	(0.074)
Log house price * bank secured (4)	0.448***	0.554**	0.815***
	(0.215)	(0.249)	(0.411)
Log GDP * bank secured (4)	–0.040	–0.025	–0.088
	(0.041)	(0.044)	(0.100)
Log transaction volume * bank secured (4)	–0.038***	–0.016***	–0.139***
	(0.004)	(0.004)	(0.010)
Log network score * bank secured (4)	–0.497***	–0.167***	–0.939***
	(0.061)	(0.059)	(0.120)
Time * credit type FE	Yes	Yes	Yes
Borrower * credit type FE	Yes	Yes	Yes
Number of observations	162,352	123,767	38,585
Adjusted *R*-squared	0.722	0.726	0.717

remains significantly more correlated with borrower-specific characteristics (transaction volumes and network score) than the two forms of bank credit.

5.7 First Wave of the Covid-19 Pandemic

Our analysis excludes the Covid-19 pandemic period characterized by some credit policies to help SMES. From April 2020, MYbank, together with 100 Chinese banks took part to the "Non-contact Loan Micro-assistance Plan" (无接触贷款助微计划).[15] While the interpretation of the results after March 2020 would be problematic because loan decisions may not be driven simply by business factors, we can analyze the effects up to March 2020, before the adoption of the micro-assistance plan.

15 Moreover, on June 1, the People Bank of China and the Ministry of Finance introduced a special purpose vehicle to channel additional funds to regional banks and big techs so that they could provide new loans to SMEs.

Table IX. Big tech credit versus bank credit controlling for demand shifts

The sample period is 2017:01–2019:04. (1) At the city-month level. (2) At the city-quarter level. Lagged one period. (3) Network score measures users' centrality in the network and is based on users' payment and funds information and social interactions. The user who has more connections gets a higher network score. (4) Dummy variable that takes the value of 1 for bank secured credit (Columns 1 and 3) or for bank unsecured cred in the first and third columns the dummy variable Bank credit takes the value of 1 for bank secured credit and 0 for big tech credit. Vice versa, in the second and fourth columns, the dummy variable Bank credit takes the value of 1 for bank unsecured credit and 0 for big tech credit. Standard errors in brackets are clustered at the city-month level. Significance level: $^{*}P<0.1$; $^{**}P<0.05$; $^{***}P<0.01$.

Explanatory variables	Dependent variable:			
	Log (credit)			
	Big tech credit versus Bank unsecured credit	Big tech credit versus Bank secured credit	Big tech credit versus Bank unsecured credit	Big tech credit versus Bank secured credit
Log transaction volume	0.037***	0.040***		
	(0.001)	(0.003)		
Log network score (3)	0.388***	0.460***		
	(0.023)	(0.046)		
Log house prices (1) * Bank credit (4)	0.048	0.431*	1.514***	1.556***
	(0.125)	(0.224)	(0.278)	(0.560)
Log GDP (2) * Bank credit (4)	0.051	−0.018	0.007	−0.068
	(0.032)	(0.042)	(0.056)	(0.109)
Log transaction volume * Bank credit (4)	−0.035***	−0.038***	−0.004	−0.055***
	(0.002)	(0.004)	(0.003)	(0.006)
Log network score (3) * Bank credit (4)	−0.358***	−0.478***	−0.186***	−0.450***
	(0.028)	(0.062)	(0.023)	(0.047)
Time * City FE	Yes	Yes	No	No
Time * Borrower FE	No	No	Yes	Yes
Borrower * credit type FE	Yes	Yes	No	No
City * credit type FE	No	No	Yes	Yes
Time * credit type FE	Yes	Yes	Yes	Yes
Number of observations	646,203	162,352	257,160	54,079
Adjusted R-squared	0.722	0.722	0.477	0.575

Table X replicates the main results of the analysis including the first wave of the Covid-19 pandemic. The specification includes interaction terms between our control variables and a Covid dummy that takes the value of 1 in February and March 2020 and 0 elsewhere. The coefficients on these interaction terms allow us to test for the presence of structural changes in the elasticities during the pandemic period.

We do not observe significant changes in the elasticities of the three different forms of credit with respect to house prices (the coefficients on the interaction term Log House Price *Covid dummy are never significant). The main results still hold: big tech credit is not correlated with house prices, while bank credit (both unsecured and secured) is positively correlated, both in normal times and during the Covid pandemic shock.

Table X. Test for structural changes during the first wave of the Covid-19 pandemic

The sample period is 2017:01–2020:03. (1) At the city-month level. (2) At the city-quarter level. Lagged one period. (3) Network score measures users' centrality in the network and is based on users' payment and funds information and social interactions. The user who has more connections gets a higher network score. (4) Covid dummy takes the value of 1 for the period 2020:02–2020:03. Standard errors in brackets are clustered at the city-month level. Significance level: $*P<0.1$; $**P<0.05$; $***P<0.01$.

Explanatory variables	Dependent variable: Log (MYbank credit) (I)	Dependent variable: Log (unsecured bank credit) (II)	Dependent variable: Log (secured bank credit) (III)
Log house price (1)	0.045	0.302***	0.495***
	(0.029)	(0.064)	(0.102)
Log GDP (2)	0.007	0.027**	–0.021
	(0.005)	(0.013)	(0.021)
Log transaction volume	0.044***	0.003***	0.005***
	(0.001)	(0.001)	(0.002)
Log network score (3)	0.470***	0.026*	–0.009
	(0.010)	(0.014)	(0.025)
Log house price * Covid dummy (1) (4)	0.005	0.023	0.023
	(0.015)	(0.024)	(0.050)
Log GDP * Covid dummy (2) (4)	–0.008	–0.043**	–0.013
	(0.011)	(0.016)	(0.032)
Log transaction volume *	–0.001	0.008**	0.013*
Covid dummy (4)	(0.003)	(0.004)	(0.008)
Log network score *	–0.023***	–0.004	0.033
Covid dummy (3) (4)	(0.008)	(0.038)	(0.052)
Time FE	Yes	Yes	Yes
Borrower FE	Yes	Yes	Yes
Number of observations	10,012,639	474,744	129,914
Adjusted *R*-squared	0.584	0.622	0.565

By contrast, the results indicate a significant reduction for the elasticity of big tech credit with respect to the network score during the first months of the pandemic, probably in connection to a change in credit risk assessment policies to support merchants who have viable business, but whose centrality measure in the big tech ecosystem decreased temporarily due to lockdown measures. Interestingly, during the first months of the pandemic bank, credit tends to be less correlated with local economic conditions (especially unsecured credit), but more reactive to firm-specific transaction volumes.

5.8 Big Tech Credit and Real Effects

As a final step, we analyze the impact of big tech credit on firms' performance in two different ways. First, we expand the analysis presented by Frost *et al.* (2019) by considering not only the impact of big tech credit use on the number of online products sold but also the impact of big tech credit offered. This should limit the endogeneity problem. Indeed, a firm that is eligible for credit and uses it (treatment group) could have better investment opportunities a priori: a higher ex-post performance could simply reflect a good ex-ante selection by the big tech credit scoring methodology. Comparing the results with those for firms that had access to credit for the first-time (but did not initially demand it) should reduce this concern.

In particular, we use the following baseline model:

$$Y_{i,t} = \alpha_i + \alpha_t + \beta_1 C_{i,\ t} + \gamma X_{i,t} + \ \varepsilon_{i,t}, \tag{5}$$

where the dependent variable $Y_{i,t}$ is the logarithm of the transaction volume for products/services sold by firm i. $C_{i,\ t}$ is the credit access dummy (or the credit use dummy), $X_{i,t}$ are borrower-specific characteristics and macroeconomic controls; α_i and α_t are, respectively, firm- and time-fixed effects; and ε_i is an error term. Standard errors are clustered at the firm level.

The results of this test are reported in the first two columns of Table XI. We find that firms that had access to the big tech credit line increased their transaction volume by 22% in the following quarter (see Column I). When we consider the use of big tech credit, the effect is smaller at 16% (see Column II).

The second test focuses on the initial offering of big tech loans. Ant Group introduced the possibility to offer MYbank credit products to QR Code merchants at the end of June 2017 and started to supply loans in August 2017. We can use this exogenous supply shock to analyze the real effects of the provision of MYbank credit on firms' transactions volumes, comparing firms with and without bank credit. We exclude August 2017 from the analysis and compare three months before (2017:05–2017:07) and three months afterward (2017:09–2017:11).

To rule out the possibility that a selection in the treatment of different firms may influence our results, we used a propensity score matching combined with a difference-in-differences type of analysis.

We first average selected firms' characteristics in the period before the launch of the new big tech loan products (pre-treatment period) and use log(transaction volume) for the pre-treatment period and average transaction volumes, gender and age of the entrepreneur, and the province j where the firm is headquartered to predict the probability of being treated. Finally, we match each firm in the treatment group with one firm in the control group that has the closest score, that is the same probability of being treated. We estimate the following Logit regression:

$$\begin{aligned}\text{Treat}_i \ = \ & \beta_1 \ln(\text{trans volume})_{i,\text{average}} + \beta_2 \ln(\text{trans volume})_{i,\text{May 2017}} \\ & + \beta_3 \ln(\text{trans volume})_{i,\text{June 2017}} + \beta_4 \ln(\text{trans volume})_{i,\text{July 2017}} + \ \beta_5 \text{Male}_i \\ & + \beta_6 \text{Age}_i + \alpha_j + \varepsilon_i,\end{aligned} \tag{6}$$

where Treat_i is a dummy that equals 1 if firm i is in the treatment group (obtain the big tech credit access in August 2017) and 0 otherwise. Matching is done using a nearest neighbor approach with a conservative Caliper equal to 0.0001. Finally, the matching is done with replacement, so that there is one match between a firm in the treatment with a firm in the control group.

Figure A1 in the Online Annex visualizes the behavior of the logarithm of transaction volumes of the two groups (treated and control) prior and after the launch of the offer of credit products by MYbank. While there is no difference between the treated and the control group until August 2017, the treatment group achieves higher levels of transactions thereafter.

Table XI. Real effects of big tech credit

Standard errors are clustered at the firm level and reported in the parentheses. *, **, and *** denote for statistical significance at 10%, 5%, and 1%, respectively. (1) The first two columns of the table report the effects of big tech credit access and use on the log of a firm transaction in the next three months [see Equation (5)]. Controls include the network score of the borrower, house prices, and GDP in the province where the firm is headquartered. (2) The last two columns of the table report the effects of the initial supply of big tech loans in August 2017 on firms' transaction volume in the following three months [see Equation (7)]. The sample has been selected by means of a propensity score matching [see Equation (6)].

	Dependent variable: Log (transaction volume)			
	Access and use of big tech credit and firm's performance (1)		Exogenous shock of credit supply (2)	
	Credit access (I)	Credit use (II)	All firms (III)	Bank access (IV)
D(Credit access)	0.225*** (0.003)			
D(Credit use)		0.164*** (0.008)		
Post∗Treat			0.143*** (0.047)	0.143*** (0.047)
Post∗Treat∗Bank access				0.021 (0.134)
Treat∗Bank access				0.014 (0.098)
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
City*Time-fixed effects	Yes	Yes	Yes	Yes
Adjusted R^2	0.454	0.453	0.588	0.588
Number of observations	9,117,297	9,117,297	111,735	111,735

We then use the following diff-in-diff model.

$$\ln(\text{transaction volume})_{it} = \beta * \text{Post}_t * \text{Treat}_i + u_i + s_{ct} + \varepsilon_{it}, \tag{7}$$

where the dependent variable is the logarithm of transaction volume for firm i and time t. The dummy Treat takes the value of 1 for those firms that received MYbank credit approval in August 2017 (only in this initial month) and zero otherwise. The variable Post takes the value of 1 after August 2017 and zero before. We control for firm-fixed effect u_i and city*-time-fixed effect s_{ct}. ε_{it} is an idiosyncratic error term.

The results in column III of Table XI show that the transaction volume increases 14.3% more for firms that had access to big tech credit (treated group) with respect to firms with similar characteristics which did not have access (control group). No significant differential effects are detected for those firms that add already access to bank credit. The coefficient on the interaction term Post∗Treat∗Bank access in the fourth column of Table XI is positive but not statistically significant.

6. Conclusions

The use of massive amounts of data by large technology firms to analyze the creditworthiness of borrower firms could replace the role of collateral in solving asymmetric information problems, with significant implications for the macroeconomy and the conduct of monetary policy.

Using a unique dataset of more than 2 million Chinese firms that received credit from both an important big tech firm (Ant Group) and traditional banks, this paper investigates how these different forms of credit correlate with local house prices, business conditions, and firm characteristics. We find that big tech credit does not correlate with house prices, but reacts strongly to firm-specific characteristics, such as transaction volumes and a network score used to calculate firm credit ratings. By contrast, both secured and unsecured bank credit react significantly to local house prices, which likely reflect useful information on the business conditions in which firms operate and on their creditworthiness. This results hold during the first wave of the Covid-19 pandemic and to a series of tests to control for endogeneity. We also show that big tech credit performs better on average ex-post in terms of defaults and of firm performance.

Our results could have important macroeconomic implications, also for the monetary transmission mechanism. They indicate that the provision of big tech credit tends to reduce the effectiveness of the "collateral channel," because the provision of credit depends less on asset price movements. At the same time, if big tech credit reacts strongly to changes in firms' transaction volume and network scores (especially for online firms), a modification in economic activity or general business conditions will be immediately reflected in credit supply. This could alter the monetary transmission mechanism and increase the effectiveness of the standard interest rate channel.

Supplementary Material

Supplementary data are available at *Review of Finance* online.

References

Adelino, M., Schoar, A., and Severino, F. (2015): House prices, collateral, and self-employment, *Journal of Financial Economics* 117, 288–306.

Aghion, P. and Bolton, P. (1997): A theory of trickle-down growth and development, *Review of Economic Studies* 64, 151–172.

Ai, H., Kiku, D., Li, R., and Tong, J. (2021): A unified model of firm dynamics with limited commitment and assortative matching, *The Review of Finance* 76, 317–356.

Albuquerque, R. and Hopenhayn, H. A. (2004): Optimal lending contracts and firm dynamics, *Review of Economic Studies* 71, 285–315.

Bahaj, S., Pinter, G., Foulis, A., and Surico, P. (2019): Employment and the collateral channel of monetary policy. Bank of England Staff Working Paper 827.

Bahaj, S., Foulis, A., and Pinter, G. (2020): Home values and firm behavior, *American Economic Review* 110, 2225–2270.

Bank for International Settlements. (2019) Big techs in finance: Opportunities and risks. BIS Annual Economic Report June.

Berg, T., Burg, V., Gombović, A., and Puri, M. (2020): On the rise of fintechs—credit scoring using digital footprints, *Review of Financial Studies* 33, 2845–2897.

Bernanke, B. (1983): Nonmonetary effects of the financial crisis in the propagation of the great depression, *American Economic Review* 73, 257–276.

Bernanke, B. and Gertler, M. (1989): Agency costs, net worth, and business fluctuations, *American Economic Review* 79, 14–31.

Bernanke, B. S., Gertler, M., and Gilchrist, S. (1999): The financial accelerator in a quantitative business cycle framework, *Handbook of Macroeconomics* 1, 1341–1393.

Besanko, D. and Thakor, A. V. (1987): Competitive equilibrium in the credit market under asymmetric information, *Journal of Economic Theory* 42, 167–182.

Bester, H. (1985): Screening vs. rationing in credit markets with imperfect information, *American Economic Review* 75, 850–855.

Boyd, J. H. and Smith, B. D. (1994): How good are standard debt contracts? Stochastic versus non-stochastic monitoring in a costly state verification environment, *Journal of Business* 67, 539–561.

Chan, Y. S. and Kanatas, G. (1985): Asymmetric valuations and the role of collateral in loan agreements, *Journal of Money, Credit and Banking* 17, 84–95.

Chaney, T., Sraer, D., and Thesmar, D. (2012): The collateral channel: how real estate shocks affect corporate investment, *American Economic Review* 102, 2381–2409.

Cloyne, J., Ferreira, C., Froemel, M., and Surico, P. (2018): Monetary policy, corporate finance and investment. NBER Working Paper 25366.

Cooley, T., Marimon, R., and Quadrini, V. (2004): Aggregate consequences of limited contract enforceability, *Journal of Political Economy* 112, 817–847.

Cornelli, G., Frost, J., Gambacorta, L., Rau, R., Wardop, R., and Ziegler, T. (2020): Fintech and big tech credit: a new database. BIS Working Papers 887.

De Roure, C., Pelizzon, L., and Tasca, P. (2016): How does P2P lending fit into the consumer credit market? Deutsche Bundesbank Discussion Papers 30.

Doerr, S. (2021): Stress tests, entrepreneurship, and innovation, *Review of Finance* 25, 1609–1637.

Favara, G. and Imbs, J. (2015): Credit supply and the price of housing, *The American Economic Review* 105, 958–992.

FSB. (2019): Evaluation of the effects of financial regulatory reforms on small and medium-sized enterprise (SME) financing. Consultative document, June.

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., and Zbinden, P. (2019): Big tech and the changing structure of financial intermediation, *Economic Policy* 34, 761–799.

Gale, D. and Hellwig, M. (1985): Incentive-compatible debt contracts: the one-period problem, *The Review of Economic Studies* 52, 647–663.

Gan, J. (2007): Collateral, debt capacity, and corporate investment: evidence from a natural experiment, *Journal of Financial Economics* 85, 709–734.

Gertler, M. and Gilchrist, S. (1994): Monetary policy, business cycles, and the behavior of small manufacturing firms, *The Quarterly Journal of Economics* 109, 309–340.

Glaeser, E. L., Huang, W., Ma, Y., and Shleifer, A. (2017): A real estate boom with Chinese characteristics, *Journal of Economic Perspectives* 31, 93–116.

Hau, H., Huang, Y., Shan, H., and Sheng, Z. (2021): Fintech credit and entrepreneurial growth. Swiss Finance Institute Research Paper 47.

Hau, H. and Ouyang, D. (2018): Capital scarcity and industrial decline: Evidence from 172 real estate booms in China. Swiss Finance Institute Research Paper 38.

He, Z. and Kondor, P. (2016): Inefficient investment waves, *Econometrica* 84, 735–780.

Huang, Y., Zhang, L., Li, Z., Qiu, H., Sun, T., and Wang, X. (2020): Fintech credit risk assessment for SMEs: Evidence from China. IMF Working Paper 193.

Holmström, B. and Tirole, J. (1997): Financial intermediation, loanable funds and the real sector, *Quarterly Journal of Economics* 112, 663–691.

Ioannidou, V., Pavanini, N., and Peng, Y. (2022): Collateral and asymmetric information in lending markets, *Journal of Financial Economics* 144, 93–121.

Jagtiani, J. and Lemieux, C. (2018a): The roles of alternative data and machine learning in fintech lending: Evidence from the lending club consumer platform. Federal Reserve Bank of Philadelphia Working Papers 15.

Jagtiani, J. and Lemieux, C. (2018b): Do fintech lenders penetrate areas that are underserved by traditional banks. Federal Reserve Bank of Philadelphia Working Papers 13.

Jeenas, P. (2018): *Monetary Policy Shocks, Financial Structure, and Firm Activity: A Panel Approach*, New York University, mimeo.

Jimenez, G., Ongena, S., Peydro, J. L., and Saurina, J. (2014): Hazardous times for monetary policy: what do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking?, *Econometrica* 82, 463–505.

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

Khwaja, A. and Mian, A. (2008): Tracing the impact of bank liquidity shocks: evidence from an emerging market, *The American Economic Review* 98, 1413–1442.

Kocherlakota, N. R. (1996): Implications of efficient risk sharing without commitment, *The Review of Economic Studies* 63, 595–609.

Lian, C. and Ma, Y. (2021): Anatomy of corporate borrowing constraints, *The Quarterly Journal of Economics* 136, 229–291.

Luohan Academy. (2019): Digital technology and inclusive growth. Hangzhou.

Lorenzoni, G. (2008): Inefficient credit booms, *The Review of Economic Studies* 75, 809–833.

Mian, A. and Sufi, A. (2011): House prices, home equity-based borrowing, and the US household leverage crisis, *The American Economic Review* 101, 2132–2156.

OECD. (2019): Financing SMEs and entrepreneurs 2019: an OECD Scoreboard. Paris.

Ottonello, P. and Winberry, T. (2020): Financial heterogeneity and the investment channel of monetary policy, *Econometrica* 88, 2473–2502.

Schmalz, M. C., Sraer, D. A., and Thesmar, D. (2017): Housing collateral and entrepreneurship, *The Journal of Finance* 72, 99–132.

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

Waxman, A., Liang, Y., Li, S., Barwick, P. J., and Zhao, M. (2020): Tightening belts to buy a home: consumption responses to rising housing prices in urban China, *Journal of Urban Economics* 115, 103190–103190.