TechFin at Ant Financial: Credit Market Completion and its Growth Effect

Harald Hau
University of Geneva, CEPR and Swiss Finance Institute[1]

Yi Huang
Graduate Institute of International and Development Studies[2]

Hongzhe Shan
University of Geneva and Swiss Finance Institute[3]

Zixia Sheng
Ant Financial Group[4]

December 19, 2017

Abstract

Ant Financial provides automated credit lines to more than a million firms trading on Alibaba's Taobao e-commerce platform. Monthly credit records show how TechFin mitigates local credit supply frictions in China's segmented credit market and extends the "frontier" of credit availability to firms with a low credit score. We use a discontinuity in the credit decision algorithm to document that a firm's credit approval and first-time online credit use boosts firm development in terms of sales and transaction growth. These findings reveal the scope of China's credit market frictions.

JEL Classification: G20, G21, O43
Keywords: TechFin, credit constraints, micro credit, entrepreneurship

[1]Geneva Financial Research Institute, 42 Bd du Pont d'Arve, 1211 Genève 4, Switzerland. E-mail: prof@haraldhau.com. Web page: http://www.haraldhau.com.
[2]Graduate Institute of International and Development Studies, Maison de la paix, Chemin Eugène-Rigot 2, 1202 Genève, Switzerland. E-mail: yi.huang@graduateinstitute.ch.
[3]Geneva Financial Research Institute, 42 Bd du Pont d'Arve, 1211 Genève 4, Switzerland. E-mail: Hongzhe.Shan@unige.ch.
[4]Director of AI Decision Science.

Acknowledgment and Diclaimer

We thank seminar participants at the Geneva International Macroeconomics and Financial Workshop, August 17-18, 2017. We thank Tony Berrada, Rajna Gibson Brandon, Stijn Claessens, Long Chen, Hongbin Gao, Giacomo De Giorgi, Christopher Malloy, Marcelo Olarreaga, Cheng Ouyang, Jun "QJ" Qian, Jean-Charles Rochet, Hong Ru, Piao Shi, Hyun Song Shin, Tao Sun, Fabio Trojani, Shang-Jin Wei and Hao Zhou for their comments. We are particularly grateful to colleagues at Ant Financial and AliResearch for their support and Sibo Liu at Hong Kong University for generously sharing his GPS data on Chinese bank branches. This research project benefited from a research grant from the Graduate Institute and a Sinergia Research Grant from the Swiss National Science Foundation (SNSF). The authors highlight that the data and analytics 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 Financial. All statements made reflect the private opinions of the authors and do not express any official position of Ant Financial and its management. The analysis was undertaken in strict observance of the Chinese law on privacy.

1 Introduction

One of the most acclaimed TechFin applications as of 2017 is the automated small firm or vendor credit offered by Ant Financial to more than a million entrepreneurs and shop owners in China's digital economy.[1] Based on the data generated by billions of transactions in Alibaba's trading platform Taobao, the financial firm Ant Financial (formerly Alipay) is able to construct automated credit ratings and provide small loans to a large number of online shops while continuously monitoring credit risk. This article provides the first comprehensive micro-economic analysis of the demand determinants of small firm credit throughout China as well as its effect on firm development.

In particular, we examine the triple hypothesis that TechFin credit by Ant Financial (i) integrates a geographically fragmented Chinese credit market characterized by important credit supply frictions for small private e-commerce firms; (ii) expands the "frontier" of credit supply to many young Taobao firms of high credit risk; and (iii) boosts the growth of such firms in the month following credit approval.

The first hypothesis focuses on online credit use by China's online entrepreneurs in the presence of credit market frictions. We measure three frictions highlighted in the previous literature on banking in China (Allen, Qian, and Qian, 2008; Qian, Strahan, and Yang, 2015). First, the overall level of banking development and loan supply in a city is captured by the aggregate *Loan/GDP* ratio and is very heterogeneous across China's city-level prefectures. A more abundant local loan supply by traditional banks should reduce the demand for online credit. Second, state-owned enterprises (SOEs) generally benefit from preferential credit access in China's banking system. A high *SOE Employment Share* (relative to total city employment) therefore shrinks the residual credit supply available to the private business sector. A third type of friction comes from a firm's geographic distance from bank branches. Based on previous banking research on the role of geography (e.g., Petersen and Rajan, 2002; Degryse and Ongena, 2005; Alessandrini, Presbitero, and Zazaro, 2009; Agarwal and Hauswald, 2010), we conjecture that a larger average distance (*AvDistance*) of a firm from the 20 closest bank branches increases search costs and information frictions in the traditional banking system. As Ant Financial's online credit dispenses with the

[1] We use the term TechFin rather than FinTech following the terminology suggested by Zetzsche *et al.* (2017): "TechFins rely on large-scale data sets [...] developed in their primary course of business and then put them to use in financial services". This is a good description of Ant Finance's business model, which is based on access to Alibaba's e-commerce transaction data.

physical presence of a bank branch and the filing of credit documents it should be most attractive to those entrepreneurs who do not have banking services in their proximity. We find evidence for all three types of credit supply frictions reflected in the online credit use: A lower local loan supply at the city level measured by the *Loan/GDP* ratio; a greater credit demand from SOEs captured by the local *SOE Employment Share;* and a larger average distance to local bank branches all have a sizeable positive impact on online credit use.

Our second hypothesis concerns the firm type benefiting most from the availability of online credit. In the presence of adverse selection (Stiglitz and Weiss, 1981) or moral hazard (Holmstrom and Tirole, 1997), Taoboa firms with a low credit score are most likely to be credit constrained. For these firms, real-time sales monitoring through the Taoboa trading platform endows Ant Financial with a potentially significant information advantage over traditional banks. We therefore conjecture that the use of online credit increases in a firm's credit risk. This second hypothesis has also strong support in the data: The extensive margin of online credit use almost triples from around 15% for the 12.5% of Taobao firms with the highest credit score to more than 40% for the 12.5% of firms with the lowest credit score. Moreover, the 25% of Taobao firms with lowest credit score show very little sensitivity in their online credit use to geographic frictions of the traditional bank branch network. We interpret this as evidence that these firms are anyway excluded from traditional bank credit so that geographic bank distance no longer matters. In other words, Ant Financial extends the frontier of credit availability to hundreds of thousands of firms with a low credit score and no access to traditional bank credit.

The third hypothesis is about the causal link between online credit approval and firm development. Here we exploit a discontinuity in the credit approval algorithm to separate similar firms with and without credit access and compare their short-term firm development. In particular, we examine the role of credit access for *Sales Growth*, *Transaction Growth,* and *Product (Variety) Growth*. Identification of these causal effects allows a better understanding of the real costs of financial frictions in terms of lower private sector growth in the world's largest economy. We find that credit approval causes a large average 18% (15.5%) jump in sales growth (transaction growth) in the month following credit approval. Two limitations of this quantitative inference need to be highlighted. First, Ant Financial's ongoing selection of new firms for credit eligibility make it difficult to define suitable control groups to identify the persistence of any growth effect. However, an 18% average growth effect is economically significant even if it is transitory. Second,

the currently available data not allow us to distinguish between sales diversion from other vendors and a genuine increase in (aggregate) consumer demand. Therefore, we cannot extrapolate our estimates to any macroeconomic prediction about the growth contribution of Ant Finance credit.

Ant Financial's business model is based on the analysis of Alibaba's e-commerce transaction data and the competitive information advantage these data provide over traditional banks. Economic models of credit rationing (Stiglitz and Weiss, 1981) and adverse selection (Holmstrom and Tirole, 1997) have highlighted the role information frictions play in the functioning of credit markets. Yet, it is generally difficult to identify distinct changes in information technology within the banking sector. Historically, banking technology has not been subject to any large disruptions in the screening and monitoring processes. But the emergence of TechFin firms like Ant Financial could represent such a disruption, characterized by significant information differences across credit providers. To what extent such new information technology extends credit availability to a larger group of otherwise credit-constrained firms has not (to our knowledge) been examined in the existing literature and represents an interesting case study in its own right.

The analysis in this paper was made possible because Ant Financial granted us access to end-of-the-month records of all e-commerce firms/vendors trading on Taobao during the period from September 2014 to July 2016. The sample of all active Taobao vendors comprises 365 million firm-months for 29.4 million firms or vendors throughout China. But many of these micro firms have only a transitory existence. Based on daily transaction data from the Taobao e-commerce platform provided by Alibaba and other financial data sources, Ant Financial subjects these firms first to a screening process and then to a second more detailed credit evaluation. For example, Taobao registered approximately 12 million active firms as of January 31, 2015. Only 1.37 million passed the first screening process and 1.08 million were granted a credit line after a second credit evaluation. During the entire sample period from September 2014 to July 2016, Ant Financial provided "Taobao credit" approval to 2.9 million firms, thereby extending online credit to a total of 28.6 million firm-months.[2] For credit-eligible Taobao firms we observe the time-averaged monthly use of the credit line and its interest rate. We also dispose of detailed information on the firm owner and firm characteristics including the firm's location, and monthly firm performance measures. We further match the firm data from Ant Financial with city-level data sourced from the *China*

[2] Ant Finance provides a variety of other types of credit to particular user groups of Alibaba's trading platforms. This paper only focuses on the Taobao vendor credit line, which is the economically most significant loan category.

City Statistical Yearbook and a detailed registry of all Chinese bank branches with their matched GPS coordinates from Google Geocoding Service.

We organize the paper follows: A short literature review is provided in Section 2, a description of the institutional setting is given in Section 3, and a discussion of the research design and the main hypotheses follows in Section 4. Section 5 is dedicated to the analysis of determinants of online credit use, and Section 6 to the causal effect of credit access on firm performance; Section 7 concludes.

2 Literature

Financial development is about extending the boundary of credit availability to a larger group of potential creditors based on better information acquisition, information processing or contracting while maintaining (or even reducing) credit risk. The TechFin model of financial intermediation seeks to extend this boundary based on new data sources and automated computer-based credit analysis. An important empirical question not yet addressed in the academic literature is to what extent such new data technologies can indeed provide credit to entrepreneurs who either do not obtain traditional bank credit or only with less favorable conditions. Issues of financial exclusion represent an important theme in the development literature.

There is a large empirical literature on the role of firm credit and its real effects. Some of the most recent contributions have used natural experiments in pursuit of better causal identification. For example, Barrot and Nanda (2017) use the acceleration of receivable payments for small government contractors after the introduction of *Quickpay* to document employment effects of macroeconomic significance. Credit access can also provide an important competitive advantage and the lack of it may increase firm survival risk. Barrot (2016) looks at the French trucking industry to show that legal restrictions on the amount of trade credit that trucking can provide to customers benefitted mostly small credit-constrained firms and lowered their default risk. By providing credit access to new entrants into the e-commerce market, Ant Financial enhances their competitiveness, as shown by jumps in sales growth, transaction growth, and product variety around the discontinuity in the credit approval algorithm.[3]

[3]This paper does not quantify the general equilibrium effects of Ant Financial's online credit, namely to estimate the percentage of firm growth that comes at the expense of sales diversion from other vendors.

This research paper relates to the literature on trade credit. Large industrial suppliers often provide their credit-constrained buyers with trade credit based on an information advantage (relative to banks) about a buyer's creditworthiness as well as potential advantages in monitoring and controlling the buyer (Petersen and Rajan, 1997). Moreover, suppliers often have an implicit equity stake in their downstream buyers. Similarly, Alibaba and Ant Financial benefit from any expansion of the e-commerce market through reduced credit frictions.

Our evidence relates to the role of credit constraints for China's growth. The role of informal lending channels and their efficiency in channeling funding to China's fast-growing private firm sector has been the subject of much debate (Tsai, 2002; Allen, Qian and Qian, 2008; Linton, 2008). Allen, Qian, and Qian (2005) argue that informal networks dispose of a screening and monitoring technology that makes the lack of access to traditional banking a lesser concern for private Chinese firms; hence their fast growth in the last two decades. However, Ayyagari, Demirgüç-Kunt, and Maksimovic (2010) find evidence to the contrary in firm surveys: private sector SMEs with access to bank credit appear to grow faster than private firms that rely on informal lending channels even after controlling for the selection bias of such a comparison. The growth effect of Taobao credit documented in this paper suggests that informal credit is a highly imperfect substitute for formal credit even for very small firms in China.

Finally, we contribute to a growing literature on the geographic segmentation of China's credit market (Dobson and Kayshap, 2006; Roach, 2006; Boyreau-Debray and Wei, 2004 and 2005; Brandt and Zhu, 2007; Dollar and Wei, 2007; Firth, Lin, Liu, and Wong, 2009). Most recently, Huang, Pagano, and Panizza (2016) provide evidence on China's credit market segmentation based on the crowding out of private investment by local government debt. Gao, Ru, Townsend and Yang (2017) provide an interesting discussion of Chinese bank deregulation aimed at reducing segmentation of the local credit market.

3 TechFin Credit in China

3.1 Alibaba and Alipay

Founded in 1999 by Jack Ma in Hangzhou (China), Alibaba counts as the most successful Chinese company. It owns three major trading platforms called Alibaba (B2B), Tmall (B2C) and Taobao

(C2C) and its commercial success in China's e-commerce sector led to a public listing in 2014 on the New York Stock Exchange. By 2016, the gross merchandise trading volume (total trading value) in the two trading platforms Tmall and Taobao exceeded 3 trillion RMB a year, which amounts to 4% of Chinese GDP.

In 2002, Alibaba began to collect data from its e-commerce platforms in pursuit of better credit information. Alipay was launched in 2004 to provide improved payment services for online transactions. In collaboration with China Construction Bank and Industrial and Commercial Bank of China from 2006 to 2010, Alipay began to grant selected loans and developed its credit rating system, databases and risk management systems. By 2010, Alipay was in a position to provide automated firm credit based on algorithms using the transaction and financial data obtained from Alibaba's online platforms. By 2014, Alipay had become the world's largest mobile and online payments platform and accounted for approximately half of all online transactions in China.[4] In 2015, Alipay was re-branded as Ant Financial Service Group. It continues to use the trading data from Alibaba's trading platform to provide automated small business credit.

Table 1, Column 1, documents the evolution of the trading volume on Taobao for the month of February in the period 2012-16. Trading grew by roughly 74% a year over this four-year period. The trading activity generates commercial data on millions of small businesses, which can be employed to alleviate the credit constraints of the most successful entrepreneurs. The formidable growth of Alibaba's online trading is often attributed to the creation of escrow accounts managed by Alipay. Securing online payments through escrow accounts represented an astute operating mode in a retail market characterized by low consumer confidence in the reliability of online counterparties.

Table 1, Columns 2-3, describe the evolution of Taobao firm loans in terms of the number of eligible firms and the number of vendors using these loans. Column 4 shows the aggregate credit volume offered, which grew by 82% a year from February 2012 to February 2016 in line with the overall growth in trading activity on Taobao. Column 6 reports the corresponding time-averaged monthly credit use, which aggregates to approximately 17% of total credit offered by Ant Financial. The aggregate Taobao firm credit volume reaches roughly 87 billion RMB in February 2016. This amounts to only 0.37% of the total micro and small firm credit supplied by China's

[4]See Bobsguide, February 12, 2014: http://www.bobsguide.com/guide/news/2014/Feb/12/alipay-surpasses-paypal-as-leading-mobile-payments-platform/

banking system.[5] From a macroeconomic perspective, Ant Financial's credit volumes are therefore small and without large repercussions for credit market equilibrium.

Yet this low aggregate credit volume may understate the economic significance of TechFin credit. First, the growth rates of TechFin credit volumes are very large—implying that its economic role is increasing. Second, Ant Finance's credit is concentrated in a very dynamic segment of China's small business sector, which gives it larger economic significance. Third, the credit might primarily go to businesses that cannot obtain credit from the traditional banking system and could therefore have a larger growth effect. A key empirical question of our analysis is therefore which vendors use the new credit technology? Taking a revealed preference argument, online credit users found the new credit technology superior to previously available sources of informal finance and traditional bank finance or were simply credit constrained.

3.2 Ant Financial's Credit Approval Process

There are two steps to Ant Financial's credit approval process. First, firms have to pass certain exclusion criteria before they undergo a second more comprehensive eligibility evaluation. Unlike traditional banks, Ant Financial can improve its credit approval process with a rich set of data sources generated by Alibaba's trading platforms. These data are used in both stages of the credit approval process.

The first step of the approval process comprises eligibility hurdles along the following five dimensions:[6]

1. **Sales Stability:** Only firms with at least a three-month of sales history qualify for credit. Alibaba employs a time series model to predict the volatility of sales for any credit candidate. A firm's up- or downstream supply chains are analyzed in terms of stability and geographic proximity.Firms that promise a sufficiently stable sales revenue in the next month are retained.

2. **Sales Authenticity and Illegal Sales:** Alibaba excludes from credit all firm with sales volumes due mainly to self-trading or reciprocal trading with a unique counterparty. Also

[5] A research team at the Central University of Finance and Economics summarized data obtained from the CBRC for their China Micro, Small and Medium Enterprise (MSME) report (Shi, 2016) and evaluated the aggreagte outstanding micro and small enterprice (MSE) loans of China's banking system at 23.46 trillion RMB.

[6] These five criteria cover most of the exclusion cases. Some additional rare exclusion cases rely on conflict of interest rules. For example, credit cannot be provided to employees of Ant Financial or their related family.

excluded are firms that violate intellectual property laws, sell fake products, or engage in other types of fraud.

3. **Logistical Service Quality:** Credit eligibility requires a certain service standard in terms of timely shipping or product quality. Here Ant Financial can draw on Alibaba's logistical data to discard unreliable firms.

4. **Customer Ratings:** The online trading platforms produce abundant data on customer satisfaction. Credit eligibility requires a minimum level of customer satisfaction.

5. **Business Network:** Alibaba analyzes a firm's commercial relationship with other businesses and their credit-worthiness. Association with fraudulent counterparties can lead to exclusion from credit.

The second step of the credit analysis is more comprehensive. It includes a linear credit scoring model that combines historical default data for firm credit with sales and financial data mostly derived from Alibaba's electronic trading platforms. A large number of variables enter the model, but the most important concern the recent sales record of a firm recorded on the trading platforms. The credit scoring model summarizes the credit evaluation in a score ranging from 400 to 600. Figure 1 shows the distribution of credit scores for the $1,372,004$ Taobao firms that had passed the first stage of the screening process as of January 31, 2015. For any given credit score, the green bars denote the number of firms with a credit approval and the red bars above represent the number of firms without credit approval. A general rule for credit eligibility is a credit score above a certain threshold. This threshold was situated at 480 for the sample considered in the second part of the analysis.

3.3 A Fuzzy Random Discontinuity in Credit Approval

The discontinuity of the credit decision at a particular (arbitrary) threshold provides an interesting statistical opportunity to explore causal effects of credit on firm performance for a large sample of firms around the discontinuity. But two important caveats complicate the analysis.

First, the credit decision of Ant Financial's algorithm depends not only on the credit score, but also on other data inputs similar to the credit exclusion criteria described above. The credit score itself is therefore only an imperfect (fuzzy) predictor of credit eligibility. The algorithm

frequently denies credit to firms above the threshold (no-show cases), but generally does not grant credit if the credit score is below the threshold of 480. As we do not have access to the full set of decision rules excluding firms above the threshold, we can only implement a so-called fuzzy random discontinuity design (FRDD).[7]

As second complication concerns data frequency. Ant Financial makes daily credit approval decisions and the approved credit line is valid for one month. But the data in this study on the credit score and other firm statistics are recorded only at the end of each month at end-of-the-month values. Firms can exceed the credit score threshold of 480 in the middle of the month and obtain a credit line (for one month), but their credit rating drops below 480 at the end of the same month. Such a firm appears to be a crossover case because it is eligible for credit even though the end-of-the-month credit rating is below the threshold value of 480. This second type of "fuzziness" is a consequence of our infrequent (monthly) credit score recording.

As we want to count as "treated" only those firms that had credit access from the end of month t to the end of month $t+1$, we assign treatment status only to those firms recorded as having credit access at the end of month t *and* at the end of month $t+1$. The same applies to firms considered "non-treated"; these had credit neither at the end of month t nor at the end of month $t+1$. Firms flipping their credit status between two consecutive months are therefore discarded from the respective firm cohort by design.[8]

4 Research Design

4.1 Data Samples

The data examined in this paper concern monthly statistics on all 29,403,374 vendors selling in Alibaba's trading platform Taobao during the period September 2014 to July 2016. The main data set and its subsamples are documented in Table 2. At the most comprehensive level denoted as Sample A, the data comprise 365 million firm-months. For a varying share of these observations,

[7] Useful introductions to the methodology include Imbens and Kalyanaraman (2012), Lee and Lemieux (2010), and McCrary (2008).

[8] Measuring the long-run growth effect of *persistent* credit access requires us to exclude an increasing number of firms both from the treatment and control group, namely all firms flipping their credit status. This rapidly errodes the statistical power of inference on long-run effects. Alternatively, we can tack the long-run sales performance of all firms with (temporary) credit access in month $t+1$ and use the estimated growth effect as a lower bound on the sales growth of firms with persistent credit access. The latter analysis is pursued in Section 6.3.

we dispose of firm credit information (Panel A), ownership and firm characteristics (Panels B and C), and data on firm performance (Panel D) recorded at the end of each month. Data on the *Credit Approval* during the calendar month are always available as well as data on most owner and firm characteristics. By contrast, information on *Log Sales*, *Sales Growth* or *Transaction Growth* is missing in more than two-thirds of the firm-months.

Our analysis mostly relies on two subsamples denoted Sample B and Sample C. Sample B consists of all Taobao firms that obtained credit in any given month and consists of 2,899,417 firms and 28,668,016 firm-months. In Sample B, we have very few data fields with missing values. Of the data fields listed in Table 2, only the migration status (a dummy for firm owners born outside the province of the firm location) is missing in 57% of all cases because the birth-place of the firm owner is unknown. As sample B comprises all firms with credit access, we use this large sample of firms to identify the determinants of online credit use throughout China.

Sample C consists of all firms with a *Credit Score* in the range 460-500. This subsample is suitable for the causal inference of credit access on firm performance. The data focus on firms with a similar (low) credit score around the discontinuity of the credit approval at the credit score of 480. This much smaller sample includes only 217,944 firm-months with no missing information on firm characteristics and firm performance. The data set is also confined to only nine months as the discontinuity at 480 was not effective during all months covered by Sample A.[9]

In Table 2, we report the number of observations and the mean value of each variable in the three samples. Taobao firms obtain a credit approval ($Credit\ Approval = 1$) in only 7.4% of all firm months. All firms with a positive credit approval decision in a particular month join Sample B. The average credit line obtained amounts to $51,800$ RMB ($8,354$ USD), of which roughly 17% is used on any day during the month. In total, eligible Taobao firms use the credit line in 6,372,984 firm-months out of 28,668,016 (or 22%). The firms in Sample C represent a much higher credit risk with an average *Credit Score* of 480.1, compared to 533.6 for Sample B. Credit approval in Sample C occurs for 25.8% of the firm-months. The average age of these firms is only 16.7 months compared to 29.1 for all Taobao vendors (Sample A) and 44.2 months for firms with credit access (Sample B).

Table 3 provides more detailed summary statistics on Sample B. The annual *Loan Rate* ranges

[9]In Sample C, we use end-of-the-month records September 2009-November 2014 and January 2015-June 2016. After June 2015, Ant Financial changed its credit approval standard and the discontinuity at 480 no longer exists. In December 2014, the 480 cutoff was also briefly suspended and we therefore exclude this month.

from 13% to 20% according to firm characteristics and credit risk. Anecdotal evidence suggests that such rates are comparable to the interest rates offered by local banks in cases where such local credit exists. By contrast, money lenders in the informal credit market are likely to charge much higher interest rates. Ant Financial does not engage in any monopolistic credit pricing that tailors the interest rate to the competitive credit supply conditions in the local market.[10] In this sense, the online credit unifies the Chinese credit market by offering a standardized credit product with interest rates conditional only on the firm credit score, firm sales and firm type. Local variations in the use of online credit therefore reveal much about geographic variations in the availability of traditional bank credit.

4.2 Owners and Firms Characteristics

While Ant Financial can monitor credit through the daily sales record of its e-commerce firms in their screening of creditors, traditional banks rely on collateral and soft information about creditors. While we cannot control for soft information available to banks, we can control for various owner characteristics, namely:

$$X^{Owner} = \{Owner\ Age,\ Male,\ Married,\ Migrant,\ Income,\ Property\}.$$

The average *Owner Age* for vendors on Taobao is 31 years. The dummy variables *Male* and *Married* describe the entrepreneur's gender and marital status. Exactly 50% of the shop owners are women and 57% of the owners are married. The dummy *Migrant* refers to a birth-place for the entrepreneur outside the province where the firm is located. Apparently, migrants are more like to become entrepreneurs: Migrants born outside the province of firm residence account for an impressive 37% of all Taobao firm owners. By comparison, the percentage of migrant workers in Chinese cities (migrating across provinces to their work location) is only 17%.[11] Ant Financial's information system also estimates a firm owner's *Income* and *Property* based on the observed trading activity and the firm location. Both variables matter as collateral to obtain traditional bank credit, but can simultaneously influence the firm owner's decision to acquire credit.

[10] Conversations with Ant Finance representatives revealed that a location-specific interest rate had not been contemplated.

[11] According to the annual survey of migrant workers from the National Bureau of Statistics, China's total labor force was 802.9 million in 2016. Rural migrants account for 17%.

The owner characteristics are complemented by the following firm characteristics:

$$X^{Firm} = \{Credit\ Score,\ Firm\ Age,\ Log\ Sales,\ Loan\ Rate,\ Firm\ Type\}.$$

The *Credit Score (CS)* represents the outcome of Ant Financial's credit risk model; *Firm Age* refers to the number of months the firm has been in existence, *Log Sales* captures firm size, and *Loan Rate* the annualized rate charged for online credit. The loan rate of a firm varies with its *Credit Score*. Importantly, Ant Financial does not discriminate across firms as a function of the local credit competition. While such a geographically uniform credit supply policy may not maximize profits, it increases the public benefit of integrating the Chinese credit market.

Finally, we use a set of dummy variables for each major type of online shop (*Firm Type*) based on the vendor categorization of Alibaba. The median monthly sales are only 12,200 RMB (1,967 USD) for all Taobao firms and 12,594 RMB (2,031 USD) for those with credit access. Table 3, Panel C also list the five most prevalent vendor types: Women's clothing accounts for 18.7% of all online shops on Taobao, followed by men's clothing with 4.7%, cosmetics with 3.8%, second-hand products with 3.6%, and women's shoes with 3.6%.

4.3 Local Credit Supply Frictions

In a geographically segmented credit market, the availability of traditional bank credit can vary, particularly for small and new businesses of high credit risk. A uniform national credit supply by Ant Financial can help to integrate an otherwise fragmented credit market. To test this hypothesis, we use proxies for the local supply distortions. The larger the credit scarcity in a particular location, the more we expect Taobao vendors to resort to online credit to mitigate these distortions.

To capture regional variation in small firm credit supply, we use the following variables:

$$X^{Local} = \{Loan/GDP,\ SOE\ Employment\ Share,\ PCGDP,\ AvDistance\}$$

The overall development of the banking system at the city-level is captured by the local loan to GDP ratio (*Loan/GDP*). A high share of loans improves the local credit supply and renders online credit more dispensable. The local credit market conditions for small business can be adversely

affected by the presence of large and privileged creditors. State-owned enterprises (SOEs) are reputed for their capital intensity, high leverage and large credit demand. We therefore include the ratio of SOE employment relative to total employment in a city (*SOE Employment Share*) as a proxy for the credit demand pull of SOEs. As financial development can also be related to the per capita GDP (*PCGDP*), we also include this city-level control in the regressor set.

We add geographical bank distance as an extra credit market friction. Using the physical GPS location of each Taobao firm, we can determine its distance from the nearest bank branches. Bank branch locations are identified using data from the Google Geocoding Service.[12] The 20 closest bank branches are the natural candidates for a loan relationship. The average geographical distance (*AvDistance*) measured in kilometers from these 20 bank branches captures "geographic frictions" of loan access. As some Taobao firms are situated in very remote locations with very low bank density, we winsorize the variable *AvDistance* at 100 km. Figure 2 plots the distribution of *AvDistance* for the 2,899,417 Taobao firms in Sample B. The average distance from the closest 20 banks is just 0.505 km for a firm at the 10th percentile, but 50.89 km for a firm at the 90th percentile of the distribution.

Geographic credit market frictions can have two distinct sources. First, a larger average geographic distance increases the search cost, in terms of time and effort, for reaching a credit-providing bank. Second, geographic distance from banks can adversely affect the quality of borrower information, as corroborated in the banking literature (e.g., Petersen and Rajan, 2002; Degryse and Ongena, 2005; Alessandrini, Presbitero, and Zazaro, 2009; Agarwal and Hauswald, 2010). We therefore predict that a larger *AvDistance* reduces *ceteris paribus* an entrepreneur's chance of obtaining bank credit. Online credit is not subject to any geographical frictions and can mitigate the geographic disadvantages of an unfavorable firm location. Accordingly, its use should positively correlate with average bank distance.

4.4 Three Hypotheses

In a geographically segmented credit market, the availability of traditional banks varies with the local credit supply conditions. The uniform national credit supply by Ant Financial can help to complete and integrate an otherwise fragmented credit market. We test this hypothesis by

[12] The translation of bank branch addresses was undertaken by Sibo Liu (Hong Kong University) using the Google Geocoding Service. We thank him for his generosity in sharing the data.

examining how variations in the use of online credit relate to local credit supply frictions. The larger the distortions and credit supply frictions in a particular location, the more we expect Taobao vendors to resort to online credit:

Hypothesis 1: Fintech Credit Compensates Market Frictions

In locally segmented credit markets characterized by credit supply frictions, we expect online credit use by Taobao firms (conditional on owner and firm characteristics) to

(i) correlate negatively with the city-level *Loan/GDP* ratio;

(ii) correlate positively with the city-level *SOE Employment Share;*

(iii) correlate positively with a firm's average distance (*AvDistance*) from the 20 closest bank branches.

Our second hypothesis explores the competitive information advantage of Ant Financial relative to traditional banks. If TechFin can source superior information on the real activity of online vendors, we expect this information advantage to translate into a sizeable extension of credit to firms with a low credit score that are (mostly) ineligible for traditional bank credit:

Hypothesis 2: Fintech Credit and Credit Risk

If Ant Financial disposes of better information about firm credit risk relative to traditional banks, we expect this competitive advantage to matter most for firms with a low credit score. Conditional on the credit supply in any location, the use of online credit should increase as firm credit scores decrease.

While the previous two hypotheses concern correlation evidence on the use of online credit, we can also conjecture a causal effect on firm development. Under conditions of constrained credit, sudden access to online credit can boost firm growth. Our third hypothesis concerns the potential real effects of Fintech credit on small firm growth:

Hypothesis 3: Fintech Credit and Firm Development

Upon access and use of automated FinTech credit from Ant Financial, Taobao firms are able to improve their short-term sales growth, transaction growth, and increase their product variety.

All three hypotheses are based on the assumption that informal credit is insufficient to overcome the distortion of China's market for small firm credit. While some authors like Allen, Qian and Qian (2005) have highlighted the role of informal credit as a driver of private sector growth, others like Ayyagari, Demirgüç-Kunt, and Maksimovic (2010) are more skeptical about the role of informal credit. Generally, informal credit comes with high interest rates, is conditional on relationship ties, and often relies on unlawful enforcement mechanisms. These are good reasons to doubt the full substitutability of informal and formal credit even for small businesses. Here we highlight that evidence for these three hypotheses is also evidence for the limited substitutability of formal and informal credit.

A second concern is that Ant Financial's market entry changes the credit supply of competing banks and our analysis does not account for their competitive response. In this context, we note that the overall credit volume of Ant Financial is still modest and amounts to only 0.37% of the total volume of small firm credit in China. In addition, it targets the "underbanked" private small business sector. Under these circumstance, the business models of most incumbent banks is not (yet) being seriously challenged. A more structural analysis of local credit markets would require detailed loan statistics for the incumbent banks, but such data are not publicly available.

5 Evidence on FinTech Credit Use

5.1 Baseline Results

In the first part of our analysis we examine which firms use online credit. As Ant Financial's credit supply policy is uniform across China's heterogeneous local credit markets, we expect firm credit use to depend on local credit conditions for alternative sources of finance. A low take-up of credit implies that firms find cheaper sources of funding either in the local banking system or through informal lending channels. Through a revealed preference argument, a high take-up of Ant Financial credit is indicative of prior credit constraints or at best access to more expensive sources of finance. At the margin, the use of Ant Financial credit might also be influenced by a "convenience factor": The credit line is flexible in the sense that firms can increase (up to the credit limit) or repay their credit on a daily base. Alternative forms of credit in the informal market may not provide such flexibility.

We denote by *Credit Use* the daily average of a firm's credit balance over the entire month. The baseline regression for the credit demand of a firm i in month t with credit access (in Sample B) follows as

$$\ln(1 + \textit{Credit Use})_{i,t} = X_{i,t}^{Local}\beta_1 + X_{i,t}^{Owner}\beta_2 + X_{i,t}^{Firm}\beta_3 + \theta_{Firm\ Type} + \mu_t + \epsilon_{i,t}, \qquad (1)$$

where we control for firm type dummies $\theta_{Firm\ Type}$ and month fixed effects μ_t.

The size of the data set and the available (inhouse) software platform impose some econometric limitations on our analysis. First, we can only run linear and logit regressions with up to 4 million observations. We therefore sample randomly 15% of all firm-months in Sample B, which produces 4,014,000 observations. Second, the software platform at Ant Financial does not include a maximum likelihood procedure to estimate limited dependent variable models. In particular, we cannot estimate a Tobin model, even if the latter might account better for the clustering of observations at the zero credit demand.

Table 4, Columns (1)-(2) provide OLS regression results for a random subsample of Sample B. All three city-level variables are statistically highly significant. The coefficient for the city-level *Loan/GDP* ratio has the expected negative sign in line with Hypothesis 1. The use of online credit is inversely related to the overall city-level loan supply by the traditional banking system. A decrease of the *Loan/GDP* ratio by two standard deviations (or -28.22) increases the online credit demand of Taobao firms by 20.8% ($= -7.36 \times -28.22/10^3$), which represents an economically large geographic variation in online credit use. This suggests that Taobao credit compensates for the underdevelopement of traditional credit across China's geographically segmented credit market.

In cities with a higher employment share in SOEs (*SOE Employment Share*), Taobao firms also use more of the online credit from Ant Financial. This conforms to the notion that SOEs are capital intensive and represent an important (and often privileged) competitor to private firm credit. An increase of city-level SOE employment by two standard deviations ($= 28.94$) increases the local online credit demand among the eligible Taobao firms by 16.6% ($= 5.74 \times 28.94/10^3$). Presumably, the SOE sector and its credit demand are crowding out private small firm credit from the traditional banking system and thus creating more demand for online credit. Without such online credit, the development of small firms in cities with a large SOE sector is more credit-constrained. Again, this suggests that Ant Financial contributes to a leveling of credit access

conditions.

Finally, a higher per capita income in a city (*PCGDP*) comes with a higher credit demand—possibly because the returns on online business investments are higher in more developed parts of the country, or because any business development is relatively more costly. However, its variation with the (log) online credit use is of small economic significance: An increase in per capita GDP by two standard deviations (= 55.49) increases credit use by a modest 10.1% (= $1.82 \times 55.49/10^3$).

In addition to the credit supply proxies *Loan/GDP* and *SOE Employment Share* at the city level, we measure more local firm-specific market frictions. The variable *AvDistance* is defined as the average distance between the vendor location and its 20 closest bank branches. It captures increasing search costs and a diminishing quality of borrower information as bank branches become more remote from the firm location. The OLS point estimate of 8.33 implies that for every increase of *AvDistance* by 10 km the online credit demand increases by approximately 8.33% (= $8.33 \times 10/10^3$). The firm at the 10% (90%) quantile of *AvDistance* features an average bank branch distance of 0.50 km (50.89 km). The (bank-) remote firm at the 90% quantile therefore has *ceteris paribus* a 41.96% (= $0.83\% \times 50.39$) higher demand for online credit compared to the firm at the 10% quantile. Taobao online credit is therefore particularly attractive for firms in bank-remote locations.

As the firms with credit access represent a selected sample of firms with lower credit risk, it is interesting to check if the selection process based on credit risk matters for the credit demand in Eq. (1). We use Sample A to run (repeatedly) a logit selection equation and calculate the (logistic) inverse Mill's ratio for all firms with credit access.[13] In Table 4, Columns (3)-(4), we include the inverse Mills ratio as an additional control variable for selection effects. The inverse Mill's ratio is statistically insignificant. Its inclusion in the regression decreases the coefficient estimate for *Loan/GDP* from −7.36 to −8.69, for the *SOE Employment Share* from 5.74 to 1.99, and for the *AvDistance* from 8.33 to 5.54. Controlling for selection effects changes the quantitative inference, but not the qualitative results.

The large economic significance of geographic market frictions captured by the variable *AvDistance* could be due to omitted city-level variables as the density of bank branches is very heterogeneous across city-level prefectures. We therefore repeat the regression in Columns (1)-(2)

[13] As the selection equation is a logit regression (instead of a probit), we use the logistic inverse Mill's ratio $\lambda = \phi(\Phi^{-1}[\Lambda])/\Lambda$, where ϕ is the normal density function, Φ^{-1} the inverse of the cumulative normal distribution, and $\Lambda = 1/(1 + e^{-X\beta})$ the predicted cumulative density for the logit.

using city fixed effects. These results, reported in Columns (5)-(6), imply a more modest point estimate of 3.59 for the variable *AvDistance,* which now captures only within-city variation in bank distance. Conditional on city fixed effects, *AvDistance* has a standard deviation of 15.93. Increasing a firm's average bank branch distance by two standard deviations therefore increases the online credit demand by 11.44% ($= 31.86 \times 3.59/10^3$).

5.2 The Role of Owner and Firm Characteristics

It is also interesting to discuss the sign and economic significance of the control variables. A larger *Owner Age* reduces online credit use. Older firm owners might be less credit-constrained, enjoy better credit access to the traditional banking system or simply prefer less investment. By contrast, male and female firm owners use online credit to a similar extent. Marital or migrant status correlates positively with credit use in the specification with city fixed effects with 7.8% and 6.7% more online credit use, respectively. The estimated owner *Income* features a very strong positive correlation with online credit use. Income of the firm owner should correlate positively with debt capacity and/or better investment opportunities and thus with the use of online credit. The existence of (real estate) *Property* of the firm owner correlates positively with online credit use if we account for city fixed effects.

The firm characteristics are statistically highly significant variables in predicting credit use. Interestingly, a higher *Credit Score* correlates with lower credit use. This evidence is consistent with Hypothesis 2, whereby the Ant Financial credit line tends to be most useful to more credit-constrained firms with a lower credit score. For the variables *Firm Age*, *Log Sales* and *Loan Rate*, we also include quadratic terms in the regression to capture potential non-linearities in their relationship to credit use. Figure 3 plots the effect of variations in three variables $X =$ *Firm Age*, *Log Sales*, *Loan Rate* on log credit use in Panels A to C, respectively. The curved line (in red) represents the estimated quadratic effect $X\widehat{\beta}_1 + X^2\widehat{\beta}_2$ on $\ln(1 + Credit\ Use)$ estimated in Table 4, Columns (5)-(6), and plotted as a function of X, whereas the histogram (in green) provides the firm distribution. For *Firm Age* in Panel A, the estimated quadratic effect peaks at 0.92 for firms with an age of 60 months. For a firm age of 110 months, the estimated quadratic effect is approximately zero, which implies that the 60-month-old firm has on average a 92% higher credit use than the 110-month-old firm. For firms aged more than five years, the use of online credit

decreases. Panel B shows that firms with larger *Log Sales* also use more online credit. A Taobao firm at the 75th percentile of the *Log Sales* distribution (*Log Sales* = 10.64) features an estimated quadratic effect of 0.000137 [= $(-216.011 \times 10.64 + 20.303 \times 10.64\hat{}2)/10^3$] compared to −0.387 [= $(-216.011 \times 8.36 + 20.303 \times 8.36^2)/10^3$] for a Taobao firm at the 25th percentile (*Log Sales* = 8.36). The latter has therefore a 38.7% (= −0.387) lower credit use. Interestingly, firms with a higher loan rate (and therefore higher credit risk) demand more online credit. A firm at the 75th percentile of the loan rate distribution (*Loan Rate* = 16.61) in Panel C has a 46% (= 3.91 − 3.45) higher online credit use compared to the firm at the 25th percentile (*Loan Rate* = 16.17). This is again consistent with Hypothesis 2 that credit-constrained firms with a low credit score find Ant Financial's online credit particularly attractive.

5.3 Local Credit Scarcity vs. Firm Credit Risk

As online credit use is conditioned by the availability of traditional banks, it is interesting to summarize the city-specific online credit demand in a so-called "heatmap". For comparison, Figure 4, Panel A, first shows the ratio of Taobao sales volume to GDP for all Chinese city-level prefectures, where dark blue denotes the highest online sales intensity. Geographic areas with a low percentage of data points are left blank (white). Panel B presents the corresponding intensity of online credit use for the same regions. Based on the size of the city fixed effects estimated in Table 4, Columns (5)-(6), we form five groups of cities with similar credit use intensity, where the dark red color denotes the 20% of city locations with the highest online credit use. For example, e-commerce firms in China's southern costal region use online credit most intensely, whereas firms in Beijing and the surrounding prefectures use it much less.

Next, we show how online credit varies with firms' credit scores conditional on local credit scarcity. Table 5 sorts all Taobao firms by the intensity of their online credit use in their city following the color scheme in Figure 4, Panel B, where Group A represents the highest online credit use (dark red) and Group E the lowest online credit use (light red). Each group proxies for a different level of online credit use intensity and thus aggregate credit scarcity. We then sort firms within each group a second time, according to their credit score, into eight credit score bins ranging from high credit risk in Column (1) to low credit risk in Column (8). Table 5, Panel A, reports for each double sorted group the percentage of Taobao firms using Ant Financial's credit

facility and Panel B reports their average (log) credit use. For all five groups, firms with lower credit scores in the left-hand columns are more likely to use online credit (Panel A) and on average use larger amounts (Panel B). For firm locations in group E with a low aggregate level of online credit use and thus a low level of credit scarcity, the use of online credit increases from 15.23% for the 12.5% of firms in bin (8) with the highest credit scores to 36.22% for the 12.5% of firms in bin (1) with the lowest credit scores. In locations with high aggregate online credit use, the percentage of firms using online credit increases even more steeply from 17.29% for firms with high credit scores in bin (8) to 43.91% for firms with a low credit scores in bin (1).

Overall, the use of online credit increases (by construction) between 2.06 and 7.90 percentage points when comparing locations of more or less intense online credit use. Yet the variations in credit use are much larger in the second dimension of sorting: The online credit use increases by 20.99 to 26.63 percentage points when comparing firms with low and high credit scores in bins (1) and (8), respectively. In accordance with Hypothesis 2, firms with low credit scores are much more likely to draw on the availability of online credit. This confirms that Ant Financial extends the frontier of credit availability to a large number of firms with a low credit score that are presumably credit-constrained in the traditional banking system. In the next section, we provide more evidence of their exclusion from traditional bank credit.

5.4 Heterogeneity: Credit Use by Firm Credit Risk

As online credit use decreases strongly in the credit score of a firm, it is insightful to compare the determinants of online credit use across firm bins sorted on the credit score. Instead of the eight bins in Table 5, we use only four bins (sorted on the firm credit score) and draw from each a random sample of $3,934,000$ observations ($= 55\%$). We then repeat the linear regressions in Table 4, Columns (1)-(2), for each sample and report the result in Table 6.

The coefficient for the *Loan/GDP* ratio is negative for all four samples sorted on the firm credit score. The effect of the aggregate credit supply on online credit use is weaker for the 25% of firms with the highest credit score. But for the variable *SOE Employment Share*, we find that its positive effect on the online credit use is strongest for Taobao firms with the highest credit rating. This can be rationalized if firms with a low credit rating are less likely to obtain any traditional bank credit and therefore do not compete with SOEs for credit. By contrast, the most creditworthy

firms ($CS \in [75\%, 100\%]$) are crowded out by SOEs credit if the local *SOE Employment Share* is large.

The same reasoning can be applied to explain why the coefficient for the average bank branch distance (*AvDistance*) in Table 6 decreases by 66% from 4.785 to 1.63 if we compare the low credit score firm sample with the high credit score firm sample. Geographic frictions matter only for firms that are in principle eligible for traditional bank credit. If firms with the 25% lowest credit scores are mostly ineligible for traditional bank credit, bank branch proximity does not matter much to them. We also check the robustness of this finding with an alternative definition of average bank branch distance based on the 10 (rather than 20) closest bank branches. In the latter case, the coefficient for *AvDistance* decreases by 41% from 5.935 for the high credit score sample to 3.488 in the low credit score sample.

Our data do not allow us to infer directly if a firm has access to traditional bank credit or informal credit. However, the decreasing role of geographic frictions in explaining online credit use is indirect evidence that traditional bank credit is inaccessible to most of the firms with low credit scores. This suggests that many of these firms have Ant Financial as their only formal credit provider.

6 TechFin Credit and Firm Performance

Next we explore if small firm credit provided by Ant Financial has a causal effect on firm development measures such as *Sales Growth*, *Transaction Growth*, or *Product (Variety) Growth.* To establish causality, we apply a fuzzy regression discontinuity design, which exploits the fact that firms passing the credit score threshold of 480 in Ant Financial's internal credit rating model substantially increase their chance of credit approval. As other (unrecorded) information also enters the automated decision process, the discontinuity design is not sharp. Approximately 6.68% of firms in the credit score range of 460 to 480 are nevertheless eligible for online credit (crossover cases), whereas 56% of firms with a credit score between 480 and 500 still do not obtain online credit (no-show cases). The existence of crossover cases can be traced to the timing of credit decisions and our belated end-of-the-month recording of the credit score. Credit approval decisions are made on a daily basis and approval becomes possible once a firm exceeds the credit threshold of 480. If the credit score declines in the days following credit approval, it is likely that we recorded

a credit score below 480 on the last day of the month (and we have a crossover case) because credit approval is valid for at least one month. Firm statistics are also recorded at the end of the month. To assure that a treated firm has credit access from the end of month t to the end of the following month $t+1$, we require that the credit approval is renewed in $t+1$. Similarly, we require that the credit approval decision for non-treated firms in the control group is negative at the end of both month t and month $t+1$.[14] This eliminates all cases of approval flipping between t and $t+1$. Also eliminated from the cohort of firms in month t are all those which might have enjoyed credit access in some month prior to month t. Thus, any treated firm in the cohort of month t does not reappear as treated in any later cohort. By contrast, non-treated firms can appear in the control sample of various firm-month cohorts as long as their credit rating is in the range 460-500 and as long as they are not approved for credit (and become treated firms). Our analysis also ignores firms with the 1% highest and 1% lowest growth rates to limit the role of data outliers in the sample.[15]

The daily reconsiderations of any firm for credit approval by Ant Financial makes it relatively difficult to discern any long-run effects of online credit. Our monthly sampling of credit approvals implies that we measure differential firm performance only over the duration of one month. While it is possible to condition on repeated credit approval and credit denial in consecutive months, such a procedure reduces the sample size considerably and is only compatible with the logic of the random discontinuity design (RDD) if the underlying credit score were a relatively static variable with little time variation. Yet, credit scores exhibit considerable monthly variation, with a standard deviation of 11.62 for firms in the credit score range 460-500.

6.1 Credit Discontinuity Effects

Figure 5 plots the percentage share of firms in Sample C that become eligible for credit as a function of their credit score. Using a bin size of one credit score, and after fitting a third order polynomial $f(CS_{i,t})$ to the (left and right side) probability distribution, we see a probability increase of credit eligibility by approximately 30% at the credit score discontinuity of 480. Table 6, Panel A, Column (1) provides a formal regression analysis using the firm-level data. The binomial credit approval

[14] Otherwise there is no synchronicity between the treatment effect and the monthly growth statistics.

[15] This implies that we ignore firms with a monthly sales growth rate above 266.2% and below -91.5%. Such extreme firm behavior is likely to be caused by factors other than credit approval.

decision *Credit Approval*$_{i,t} \in \{0, 1\}$ is explained by the credit eligibility proxy

$$IV_{i,t} = \begin{cases} 1 & \text{if} \quad 480 \leq CS_{i,t} < 500 \\ 0 & \text{if} \quad 460 < CS_{i,t} < 480 \end{cases} .$$

The coefficient of 0.289 identifies the increase in approval probability by passing the threshold. It is statistically highly significant with a t-value of 63 and therefore constitutes a strong instrument.

Next, we demonstrate the effect of credit approval discontinuity on firm *Sales Growth*. We sort firms with credit scores (CS) in the range $[460, 500]$ into 40 bins of similar credit scores where each bin is one credit score point wide. Figure 6, Panel A, draws the average sales growth for firms in each credit score bin as red dots for the 20 bins below the discontinuity threshold of 480 and as green dots for the 20 bins above the threshold. The average sales growth of firms in bins just above the threshold of 480 is almost 5% higher than for those in credit score bins just below. Given that 29% more firms obtain credit above the discontinuity, the imputed incremental sales growth effect for treated firms is 17.2% (= 0.05/0.29) on average. The first full calendar month after credit approval is therefore characterized by an economically significant increase in online sales. Panel B provides similar evidence for the average transaction growth rate, which jumps by approximately 4% for firms just above the credit threshold of 480.

We note that the fitted polynomials (of third order) in Figure 6 shows a strong increase of the average *Sales Growth* for credit scores below the threshold, but much less so for credit scores above the threshold. As shown in Figure 5, many firms with an end-of-the-month credit score below 480 are eligible for online credit because they passed the threshold of 480 at some (for us unobservable) day within the month, but then recorded a credit score below 480 on the last day of the month. By conditioning on end-of-the-month credit scores, we generate crossover values, which reduce the magnitude of the measured discontinuity and at the same time increase the slope of the fitted polynomial below the threshold value. Visual inspection also suggests that reducing the width of the measurement window from 40 to 20 or 10 bins does not much influence the identified discontinuity.

Table 6, Panel A, Column (2) provides the reduced form regression corresponding to Figure 6, where the percentage sales growth (from month t to $t + 1$) is directly related to the instrument

$IV_{i,t}$ in the linear regression

$$Sales\ Growth_{i,t+1} = \delta_0 + \delta_1 IV_{i,t} + f(CS_{i,t}) + \epsilon_{i,t},$$

where $f(CS_{i,t})$ denotes the third-order polynomial of the credit score $CS_{i,t}$ of firm i in cohort t. The OLS coefficient δ_1 identifies the incremental average sales growth increase for firms above the threshold as 5.2%. The imputed incremental average sales growth effect for the 29% firms gaining credit approval therefore amounts to 18% (= 0.052/0.29). The actual use of this online credit for firms with first-time approval is higher at 48% than for the average take-up rate among credit-eligible firms given by 22%. The partial use of online credit implies that truly credit-constrained firms (which use the online credit in the month after their approval) experience an even larger average sales growth effect.

Table 6, Panel A, Columns (3)-(5) report the corresponding estimates of growth in the number of online transactions (*Transaction Growth*) and the variety of products offered (*Product Growth*). The estimated incremental *Transaction Growth* due to credit access is slightly smaller at 15.5% (= 0.045/0.29) than the incremental *Sales Growth*—indicating that the value per transaction increases after online credit approval. The product variety of the online vendor also increases in the month following credit approval, even though the growth effect here is much smaller at 3.8% (= 0.011/0.29).

6.2 Credit Use as the Growth Channel

The reduced form regressions in the previous section do not identify credit use as the channel through which credit approval translates into higher sales growth. Theoretically, a credit line can change the investment behavior of the online vendor even if the credit is not used. For example, precautionary motives about liquidity might deter investment and the credit line liberates liquidity for investment even if it is not used. Notwithstanding such secondary channels of growth in the absence of credit use, we generally expect that the take-up of online credit is directly linked to the observed sales growth effect. This motivates the 2SLS regression

$$Sales\ Growth_{i,t+1} = \beta_0 + \beta_1 \ln(1 + Credit\ Use)_{i,t+1} + f(CS_{i,t}) + \eta_{i,t},$$

where we use the credit eligibility proxy $IV_{i,t}$ as an instrument for log credit use and $f(.)$ again denotes the third-order polynomial function in the credit score $CS_{i,t}$ of firm i at the end of month t.

Table 3, Panel B, Columns (5) and (6)-(8) report the regression results for the first and second stage of the regression, respectively. The average log credit use $[\ln(1 + Credit\ Use)]$ is 15.9 percentage points higher for firms above the threshold than below, according to Column (5) in Panel B. The 2SLS estimate of $\beta_1 = 0.292$ for *Sales Growth* in Column (6) implies that an increase of 15.9 percentage points in credit use increases sales growth for firms above the threshold by an incremental 4.64% ($= 0.292 \times 0.159$). This estimate is reasonably close to the 5.2% obtained in the reduced form estimation in Column (2). Alternatively, we can directly quantify the growth effect for firms using Ant Financial's credit facility for the first time: Their average credit use is $9,659$ RMB relative to $2,598$ RMB for the average firm in Sample C, which implies a relative increase in growth of 38% $[= 0.292 \times (\ln(1 + 9,659) - \ln(1 + 2,598))]$ relative to firms with the average credit use.[16]

The 2SLS estimate of 0.204 for transaction growth in Column (7) implies an average *Transaction Growth* of 3.24% for firms above the credit score threshold. The average incremental transaction growth for newly credit-eligible firms follows as 16.46%. Overall, credit approval by Ant Financial provides an economically significant boost to the economic development of small e-commerce firms.

6.3 Long-Term Growth Effects

Ant Financial's continuous reevaluation of firms' online credit make it difficult to identify causal long-term growth effects of credit access. Credit eligible firms in month t may no longer be eligible in later months, and firms without credit access may become credit eligible. This means that the treatment and control group of month t become more similar in their share of firms with credit access as time passes. Panel A in Figure 7 shows how treated firms (with initial credit) loose credit access and non-treated firms (no initial credit) gain credit access for the original sampling band width ($CS \in [460, 500]$). After 6 months, more than 50% of all treated firms have lost their

[16]This is again close to the reduced form estimate in Section 6.1 of 17.2% incremental sales growth for all firms approved for online credit, which scaled by the 48% of credit take-up among credit eligible firms implies an incremental sales growth rate of 35.8% ($= 0.172/0.48$) for firms using the credit facility.

credit access, whereas more than 15% of non-treated firms have acquired online credit. Panel B depicts the analogous graph for a narrower band width of credit scores ($CS \in [470, 490]$). A narrower bank width does not alter the limited persistence of online credit access or exclusion around the discontinuity threshold at 480. Generally, estimates of the long-term growth effect conditional on a short-term credit approval for month $t+1$ should undervalue the growth benefits of uninterrupted credit access over many months. Thus, we can consider these estimates as a lower bound on the sales growth effects of persistent credit access relative to persistent exclusion.

The lower two panels in Figure 7 depict the 2SLS estimate for the effect of credit use [$\ln(1 + Credit\ Use)$] in month $t+1$ on firm sales growth from the end of month t to end of month $t+k$, where $k = 1, 2, ...6$. The point estimates in Panels C and D for the large and small sampling band are similar and roughly constant in the time horizon k. This implies that the sales growth effect is persistent beyond the first (full) month after the credit approval and does not revert. The smaller sampling band width in Figure 7, Panel D, produces a larger initial point estimate of 0.388 for $k = 1$ (compared to 0.292 in Panel C). However, the confidence intervals are larger under the narrow band width. Given the transitory nature of credit access and exclusion in the sample, it is likely that the sales growth gap between firms with persistent credit access and persistent exclusion from credit is in fact much larger. The sales growth estimates for $k \geq 2$ in Figure 7 represent only a lower bound for this comparison.

Finally, we highlight two further limitations of our analysis. First, our limited data access does not allow us to identify which percentage of the firm growth is retail demand diversion from other vendors as opposed to aggregate demand creation. This makes it difficult to estimate the precise macroeconomic growth contribution of Ant Finance's new credit technology. Second, the reported growth effects of improved credit access pertain to the e-commerce retail sector. But credit market distortions are likely to pervade all sectors of the Chinese economy, and many of those sectors are more capital intensive than the retail sector. Improved credit technologies could potentially be even more valuable for capital intensive sectors.

7 Conclusion

This paper examines the triple hypothesis that TechFin credit by Ant Financial (i) helps to complete and integrate a geographically segmented Chinese credit market characterized by important

credit supply frictions for small private e-commerce firms; (ii) makes the biggest contribution in overcoming these market frictions for Taobao firms with a low credit score (and often not served by traditional banks); and (iii) fosters sales and transaction growth of such firms in the month following credit approval.

Based on Ant Financial's credit approval records for 365 million firm-months and positive credit approvals for 28.6 million firm-months we find that the use of online credit systematically relates to local credit supply frictions. City-level measures of aggregate credit supply like the *Loan/GDP* ratio correlate *negatively* with the use of online credit. By contrast, city-level proxies of local credit market distortions like the local *SOE Employment Share* correlate *positively* with the use of online credit by Taobao vendors, after controlling for owner and firm characteristics. We also find that the geographic proximity to local bank branches matters for online credit use. The larger average distance from the 20 closest bank branches (*AvDistance*) predicts a more extensive and intensive use of online credit. This finding reflects search and information frictions related to geographic distance inherent in traditional bank credit but obsolete for online credit.

Our analysis also shows that the demand for online credit is largest for Taobao firms with a low credit score. Their online credit use is also largely insensitive to credit supply substitution effects coming from local SOE firms and to geographic frictions related to the proximity of traditional bank branches. We interpret this as evidence that these firms are mostly cut off from bank credit anyway and can only obtain online credit from Ant Financial. Its new information technology therefore extends credit access to hundreds of thousands of small firms beyond the credit "frontier" of the traditional banking system.

In the last part of our analysis, we show that TechFin credit access also has important consequences for firm growth. By exploiting a discontinuity in the probability of credit approval at a particular threshold value of the internal credit score, we provide evidence that a positive credit approval implies an economically significant development boost in the first month of credit access measured by higher *Sales Growth*, *Transaction Growth*, and *Product (Variety) Growth*. On average, Taobao firms acquiring online credit lines increase their sales growth (transaction growth) by an average of 18% (15.5%) in the month following credit eligibility. Long-run estimates suggests that the sales growth effect is persistent for at least 6 months, and could be larger for firm which enjoy uninterrupted online credit access.

These results reveal the growth constraints on small firm development resulting from credit

market frictions in China. An expansion of Ant Financial's aggregate credit volume can help to equalize growth prospects of small e-commerce firms by creating more homogeneous credit conditions, which will contribute to China's private sector growth. While our results on the growth effects of credit access pertain to the e-commerce segment of China's retail sector, it is plausible that the real costs of China's credit market frictions are quantitatively larger in more capital intensive sectors.

References

[1] Agarwal, S., and R. Hauswald. 2010. Distance and Private Information in Lending. *The Review of Financial Studies* 23(7): 2757-2788.

[2] Alessandrini, P., A. F. Presbitero, and A. Zazzaro. 2009. Banks, Distances and Firms' Financing Constraints. *Review of Finance* 13(2): 261-307.

[3] Allen, F., J. Qian, and M. Qian. 2005. Law, Financial, and Economic Growth in China. *Journal of Financial Economics* 77:57-116.

[4] Allen, F., J. Qian, and M. Qian. 2008. China's Financial System: Past, Present and Future, in Brandt, L., and T. Rawski (eds.), *China's Great Economic Transformation*, Cambridge University Press.

[5] Ayyagari, M., A. Demirgüç-Kunt, and V. Maksimovic. 2010. Formal versus Informal Financial: Evidence from China. *Review of Financial Studies* 23(8): 3048-3097.

[6] Barberis, J., and D.W. Arner, 2016. "E-commerce microcredit In China: From Shadow Banking to P2P Lending," In Tasca, P. *et al.* (Eds.), *Banking Beyond Banks and Money: A Guide to Banking Services in the Twenty-First Century*, Cham: Springer.

[7] Barrot, J.-N. 2016. Trade Credit and Industry Dynamics: Evidence from Trucking Firms, *Journal of Financial* 71(5): 1975-2015.

[8] Barrot, J.-N., and R. Nanda. 2017. Can Paying Firms More Quickly Affect Aggregate Employment? MIT working paper.

[9] Beck, T., A. Demirgüç-Kunt, and V. Maksimovic. 2005. Financial and Legal Constraints to Growth: Does Firm Size Matter? *The Journal of Financial* 60: 137-177.

[10] Bekaert, G., C. R. Harvey, and C. Lundblad. 2005. Does Financial Liberalization Spur Growth? *Journal of Financial Economics* 77: 3-55.

[11] Boyreau-Debray, G., and S.-J. Wei. 2005. Pitfalls of a State-Dominated Financial System: The Case of China? NBER Working Paper No. 11214.

[12] Brandt, L., and X. Zhu. 2007. China's Banking Sector and Economic Growth. In: *China's Financial Transition at a Crossroads*, edited by Charles W. Calormiris. Columbia University Press.

[13] Degryse, H., and S. Ongena. 2005. Distance, Lending Relationships, and Competition. *Journal of Finance* 60(1): 231-266.

[14] Demirgüç-Kunt, A., and V. Maksimovic. 1998. Law, Financial, and Firm Growth. *The Journal of Financial* 53: 2107-2137.

[15] Dobson, W., and A. Kashyap. 2006. The Contradiction in China's Gradualist Banking Reforms. *Brookings Papers on Economic Activity* 37: 103-162.

[16] Dollar, D., and S.-J. Wei. 2007. Das (Wasted) Kapital: Firm Ownership and Investment Efficiency in China. NBER Working Paper No. 13103.

[17] Firth, M., C. Lin, P. Liu, and S. Wong. 2009. Inside the Black Box: Bank Credit Allocation in China's Private Sector. *Journal of Banking and Finance* 33: 1144-1155.

[18] Gao, H., H. Ru, R. M. Townsend, and X. Yang. 2017. Rise of Bank Competition: Evidence from Banking Deregulation in China. Working paper.

[19] Holmstrom, B., and J. Tirole. 1997. Financial Intermediation, Loanable Funds, and the Real Sector, *Quarterly Journal of Economics* 112 (3): 663-691.

[20] Huang, Y., M. Pagano, and U. Panizza, 2016. Public Debt and Private Firm Funding: Evidence from Chinese Cities, CEPR DP 11489.

[21] Imbens, G., and K. Kalyanaraman, 2012. Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *Review of Economic Studies* 79: 933-959.

[22] Linton, K. C. 2008. Access to Capital in China: Competitive Conditions for Foreign and Domestic Firms. *Journal of International Commerce and Economics* 1: 27-49.

[23] Lee, D. S., and T. Lemieux. 2010. Regression Discontinuity Designs in Economics. *Journal of Economic Literature* 48: 281-355.

[24] McCrary, J. 2008. Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics* 142: 698-714.

[25] Petersen, M. A., and R. G. Rajan. 1997. Trade Credit: Theories and Evidence. *The Review of Financial Studies* 10(3): 661-691.

[26] Qian, J., P. E. Strahan, and Z. Yang. 2015. The Impact of Incentives and Communication Costs on Information Production and Use: Evidence from Bank Lending. *Journal of Finance* 70(4): 1457-1493.

[27] Raddatz, C. 2006. Liquidity Needs and Vulnerability to Financial Underdevelopment. *Journal of Financial Economics* 80: 677-722.

[28] Roach, S. 2009. China's Great Contradiction. In: *The Next Asia*, edited by Stephen Roach. John Wiley and Sons.

[29] Shi, J. P. 2016. China MSME Finance report. Mintai Institute of Finance and Banking and Central University of Finance and Economics, available at DOI: 10.13140/RG.2.2.21772.10880

[30] Stiglitz, J. E. and A. Weiss. 1981. Credit Rationing in Markets with Imperfect Information. *American Economic Review*, 71(3): 393–410.

[31] Tsai, K. 2002. *Black-Alley Banking: Private Entrepreneurs in China*. Ithaca, NY: Cornell University Press.

[32] Zetzsche, D., R. P. Buckley, D. W. Arner, and J. N. Barberis. 2017. From Fintech to Techfin: the Regulatory Challenges of Data-Driven Finance, Working Paper Series No. 6, EBI.

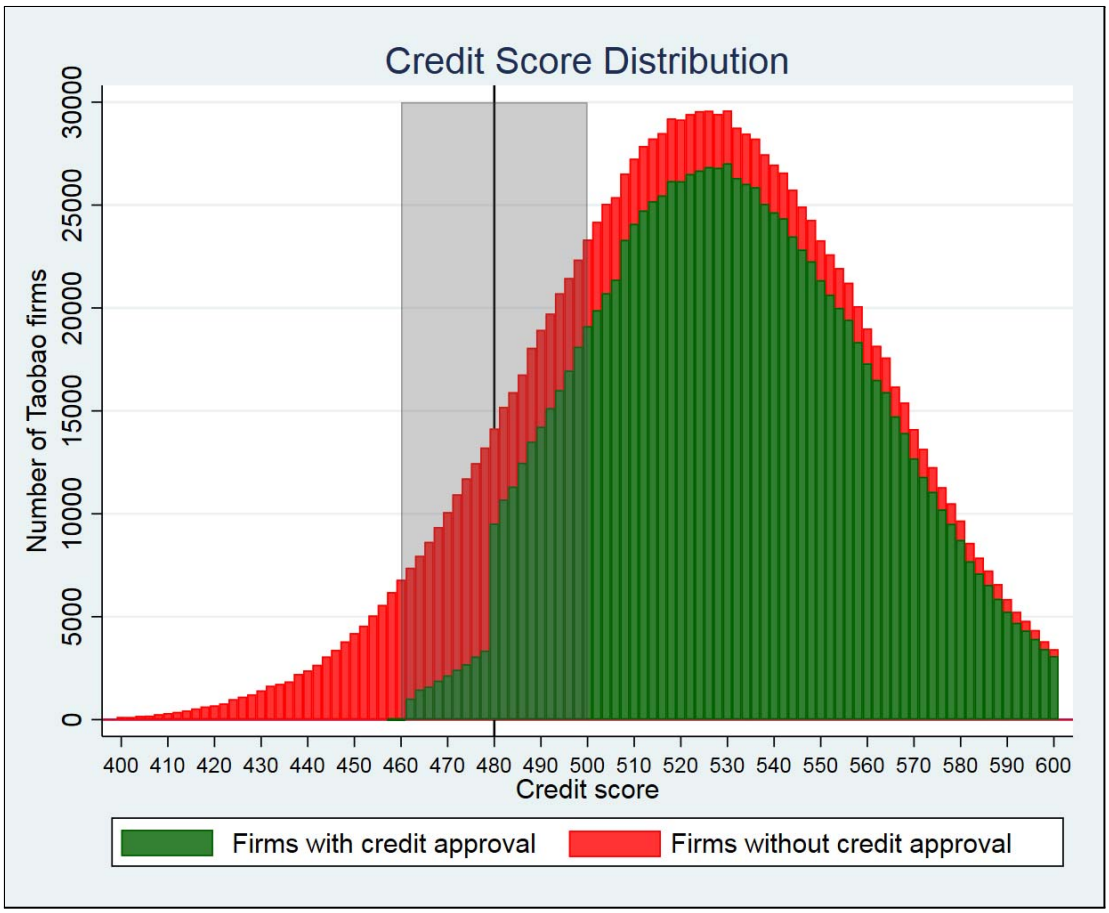


Figure 1: We plot the distribution of credit scores for 1.37 million Taobao firms that had qualified for the second step of the credit evaluation process as of January 31, 2015. The green bars mark the 79% of Taoboa firms with credit approval and the (upper) red bars represent the 21% without credit approval. The grey shaded region marks the sampling interval for the fuzzy random discontinuity design (RDD) with the discontinuity at the credit score of 480.

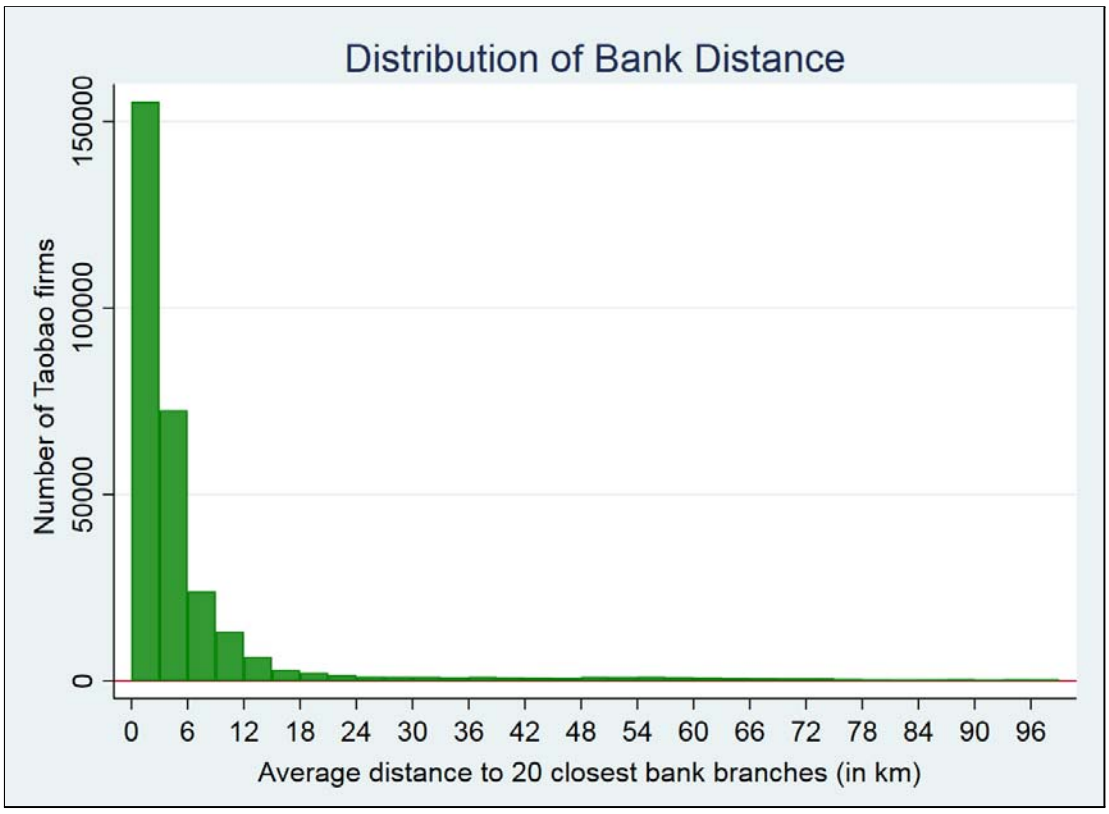


Figure 2: We plot the distribution of average distance from the 20 closest bank branches (*AvDistance*) for the 2,899,417 Taobao firms in Sample B. The bin size of the histogram is 3 kilometers.

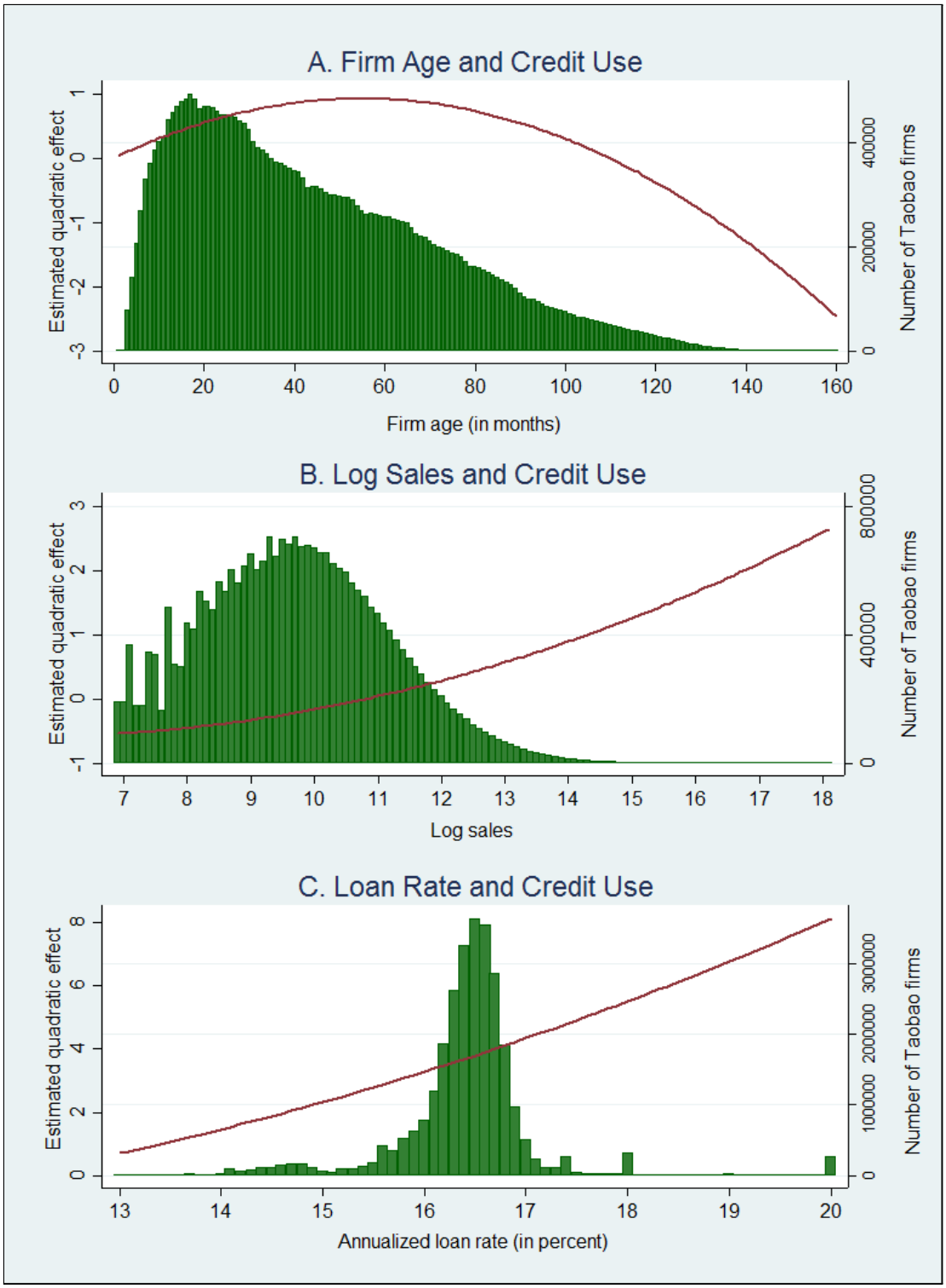


Figure 3: We plot the effect of variations in three variables, X = *Firm Age*, *Log Sales*, *Loan Rate*, on log credit use. The curved line (in red) represents the estimated quadratic effect $X\widehat{\beta}_1 + X^2\widehat{\beta}_2$ on $\ln(1 + Credit\ Use)$ estimated in Table 4, Columns (5)-(6). The bar graph (in green) provides the respective firm distribution for each variable.

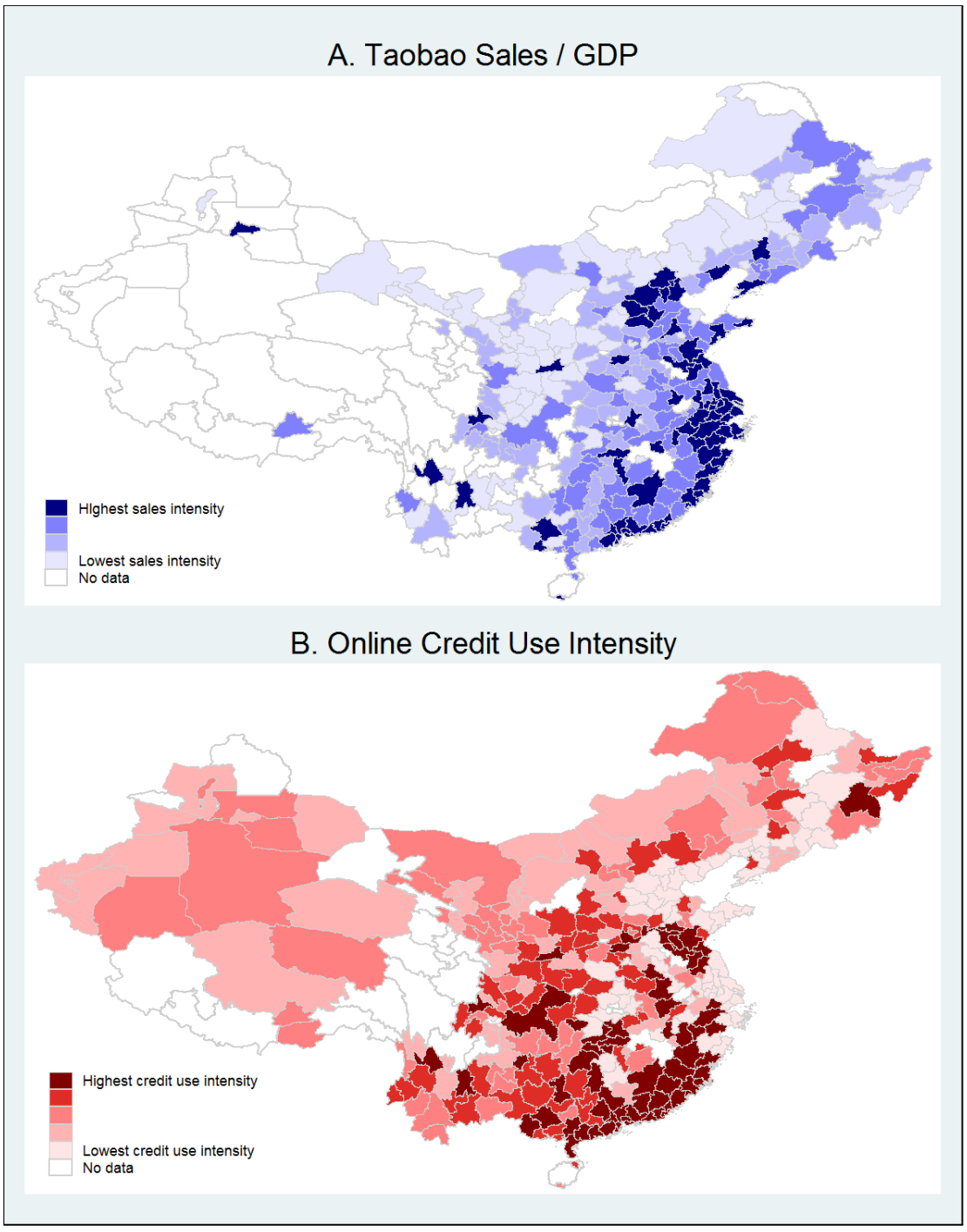


Figure 4: Panel A shows the ratio of Taobao sales volume relative to GDP for all Chinese city-level entities. Dark blue denotes the highest online sales intensity. Geographic areas with a low percentage of data points are left blank (white color). Panel B depicts the corresponding intensity of online credit use for the same regions grouped into five categories.

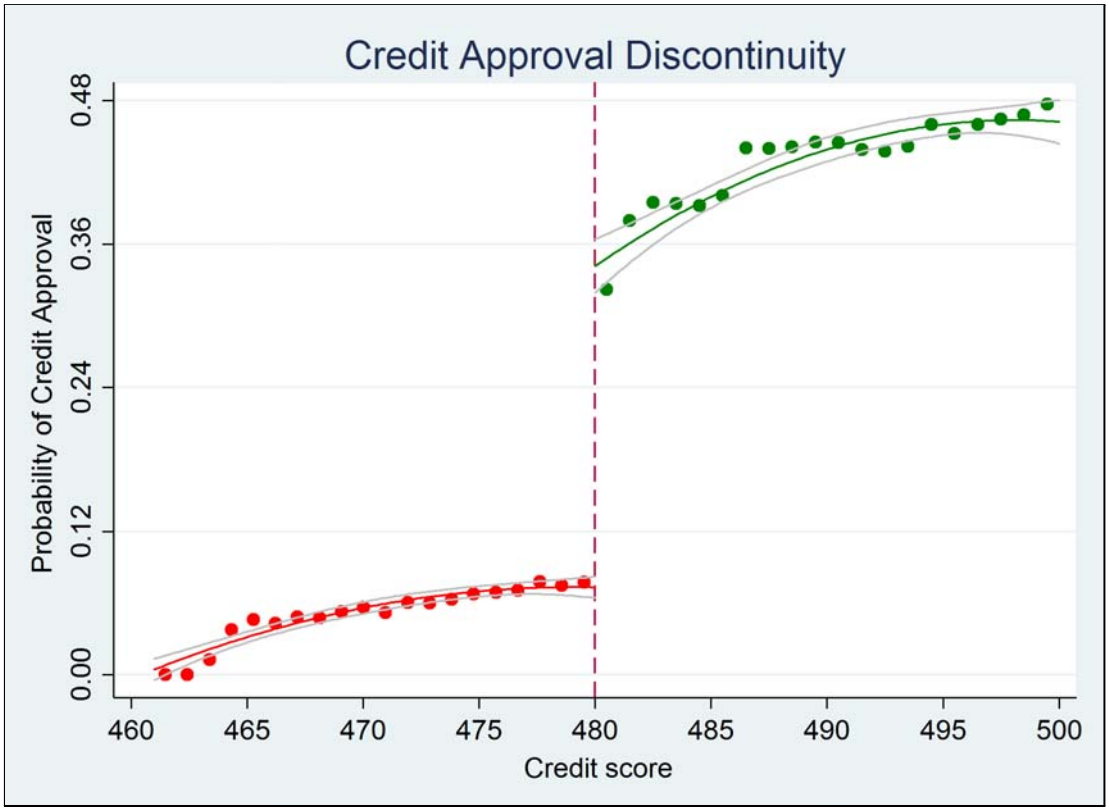


Figure 5: We plot the percentage share of firms in Sample C that become credit eligible as a function of their credit score. The bin size is one credit score unit and the curved line a third-order polynomial $f(CS_{i,t})$ fitted to the probability distribution on both sides of the discontinuity at the credit score of 480.

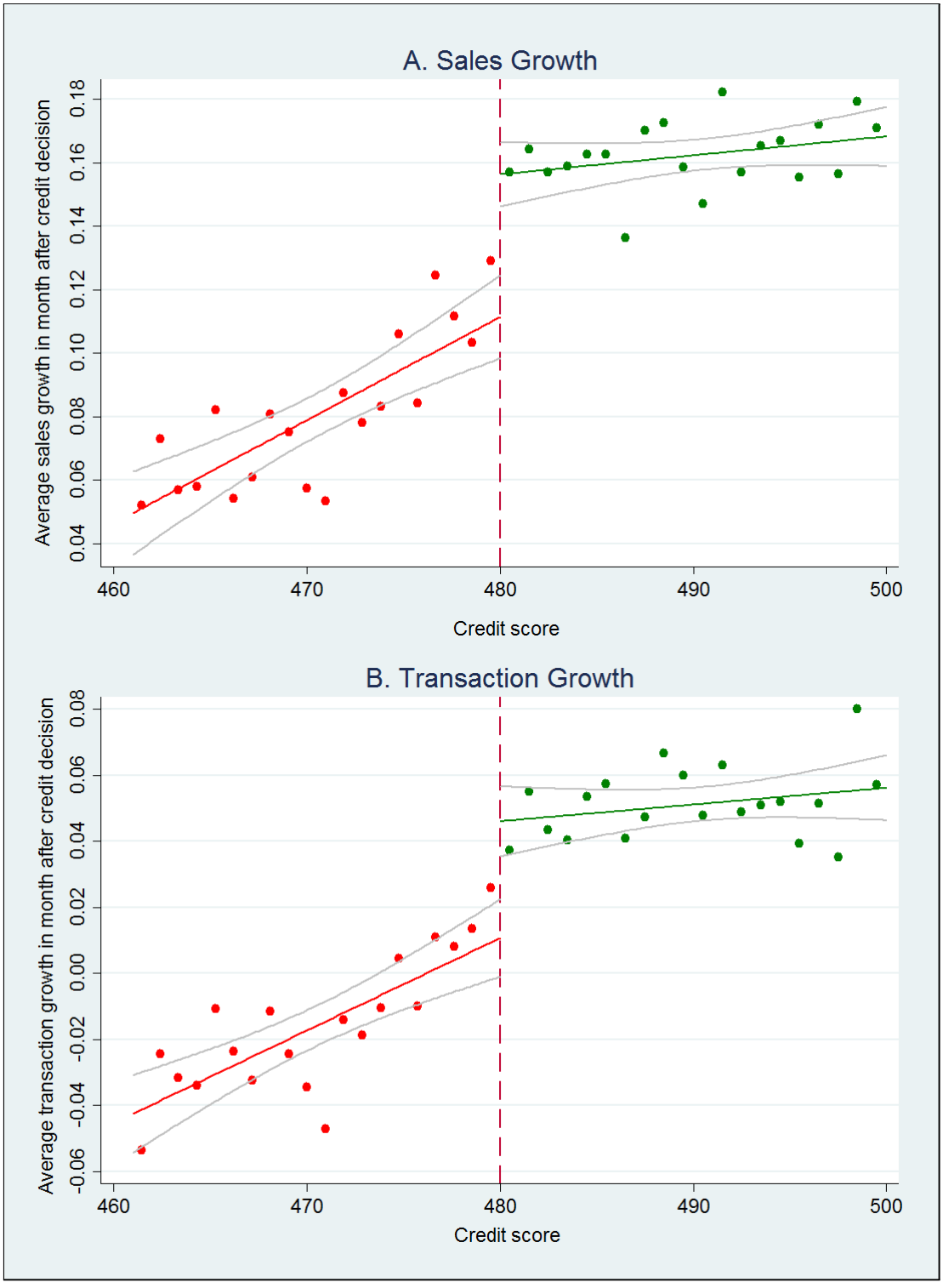


Figure 6: Panels A and B depict average sales growth and average transaction growth, respectively, for firms in each credit score bin. Red dots mark the averages for the 20 bins below the discontinuity threshold at 480 and green dots for the 20 bins above the threshold.

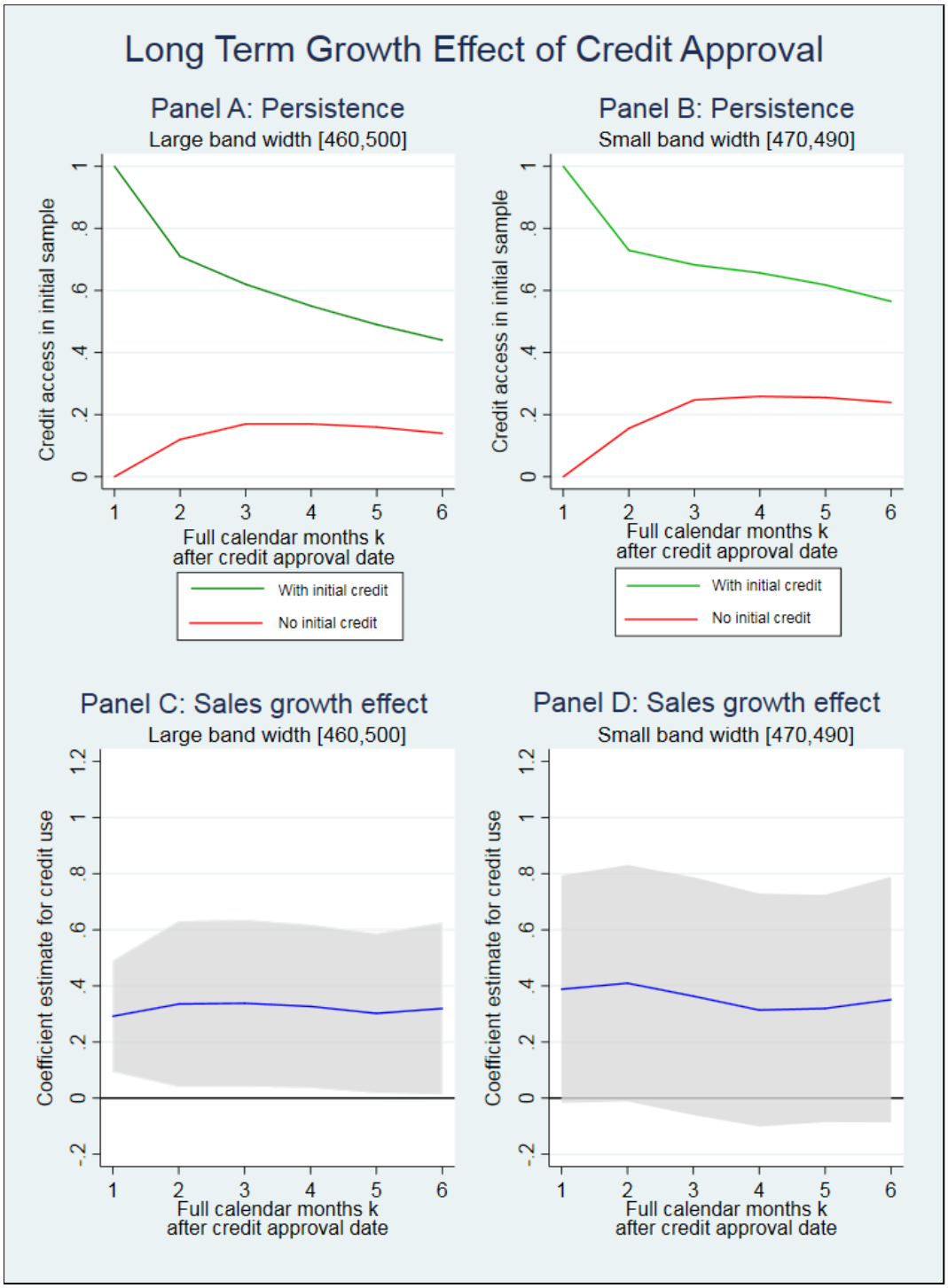


Figure 7: For the large and small sampling band we identify the treated (= with initial credit) and non-treated (= no initial credit) firms in period $t+1$ and track in Panels A and B, respectively, the persistence of their credit access and credit exclusion in the consecutive months $k = 2, 3, 4, 5, 6$. Panels C and D show the respective 2SLS estimates for the relationship between sales growth from the end of month t to $t+k$ and credit use measured as $\log(1+Credit\ Use)$. The grey area around the point estimates (bleu line) corresponds to a confidence interval of ± 2 standard deviations.

Table 1: Evolution of Firm Credit

Column (1) reports the evolution of trading volume (from February 2012 to February 2016) in the e-commerce trading platform Taobao. The trading activity provides the information for the daily allocation of automated credit lines to the firms/vendors on Taobao. Columns (2) and (3) report the number of firms eligible for credit and those using at least some of the online credit, respectively. The total credit volume offered is reported in Column (4) and the outstanding credit (credit used) in Column (5). Trading and credit volume are expressed in billions of RMB.

	Taobao	Firm Credit			
	Trading Volume	Number of Firms		Credit Volume	
		Eligible	Use Credit	Offered	Outstanding
	(1)	(2)	(3)	(4)	(5)
Feb 2012	824	95,645	11,842	47.2	5.1
Feb 2013	1173	310,946	33,968	116.7	13.8
Feb 2014	1597	751,920	152,685	263.0	44.2
Feb 2015	1877	1,103,183	231,512	429.2	75.1
Feb 2016	2202	883,294	310,486	516.9	86.7

Table 2: Summary Statistics for Samples

We report the number of available firm-month observations and the mean values for various firm statistics in three different samples covering the period September 2014 to July 2016. Sample A includes all Taobao e-commerce firms considered by Ant Finance prior to the two steps of the credit approval process. Sample B reports on the subsample of firms that obtain credit access (*Credit Approval = 1*). Sample C is used for a random discontinuity design and reports on all firms that passed the first step of the credit approval process and obtained a credit score in the range [460, 500] at the end of the month. Sample C excludes the month in which the credit discontinuity at the credit score of 480 was not effective. Panel A lists the available variables on firm credit information, Panel B lists the variables concerning the firm owner, Panel C lists the available variables for firm characteristics, and Panel D for firm performance. For details on the definition of the variables see Table 5.

		Sample A Taobao firms		Sample B Firms with credit access		Sample C Firms near credit threshold	
		Observations	Mean	Observations	Mean	Observations	Mean
		(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Online credit information							
Credit Approval	(0/1)	365,027,776	0.074	28,668,016	1	217,944	0.258
Credit Line	Approval =1	28,668,016	51,800	28,668,016	51,800	56,303	14,438
Credit Use/ Credit Line	Approval =1	28,668,016	0.372	28,668,016	0.372	56,303	0.468
Ln(1+Credit Use)	Approval =1	26,980,677	2.097	28,668,016	2.097	56,303	5.077
Ln(1+Credit Use)	*Credit Use* >0	6,372,984	9.434	6,372,984	9.434	14,710	8.490
Panel B: Owner characteristics							
Owner Age	Years	359,652,035	31.6	27,852,659	34.3	217,944	38.0
Male	(0/1)	360,429,191	0.500	27,856,705	0.528	217,944	0.548
Married	(0/1)	328,916,868	0.571	26,819,021	0.680	202,388	0.516
Migration	(0/1)	109,746,867	0.372	12,252,990	0.448	41,668	0.485
Income	RMB	362,731,541	5806	28,515,630	6446	217,944	5751
Property	(0/1)	355,048,759	0.552	27,388,906	0.541	217,944	0.361
Panel C: Firm characteristics							
Firm Age	Months	364,706,368	29.1	26,980,677	44.2	217,944	16.7
Log Sales		111,953,411	7.807	26,980,677	9.178	217,944	9.508
Credit Score (CS)		364,969,202	513.4	26,980,677	533.6	217,944	480.1
Firm types (5 most frequent)							
Women's Cloth	(0/1)	365,027,776	0.187	26,980,677	0.126	217,944	0.278
Men's cloth	(0/1)	365,027,776	0.047	26,980,677	0.044	217,944	0.109
Cosmetic	(0/1)	365,027,776	0.038	26,980,677	0.038	217,944	0.082
Second-hand prod.	(0/1)	365,027,776	0.036	26,980,677	0.038	217,944	0.032
Women's shoes	(0/1)	365,027,776	0.036	26,980,677	0.037	217,944	0.018
Panel D: Firm performance							
Sales Growth	Monthly	98,921,387	1.740	26,980,677	0.883	217,944	1.683
Transaction Growth	Monthly	105,826,924	1.873	26,980,677	0.923	217,944	1.892
Product Growth	Monthly	301,163,086	0.994	26,980,677	0.473	217,944	0.678

Table 3: Summary Statistics for Firms with Credit Access

We report summary statistics for all firm-months in which Chinese e-commerce enterpreneurs had credit approval from Ant Finance during the period September 2014 to July 2016. Panel A summarizes firm credit information,, Panel B provides ownerhsip data, Panel C reports data on firm characteristics, Panel D provides firm performance, and Panel E summarizes local credit market conditions. For details on the variable definitions see Tables 5.

		Observations (1)	Mean (2)	STD (3)	Skewness (4)	P10 (5)	P50 (6)	P90 (7)
Panel A: Online credit information								
Credit Line	RMB	28,668,016	51,800	147,575	5.559	10,000	12,000	100,000
Credit Use/Credit Line		28,668,016	0.372	0.597	26.606	0	0	0.833
Ln(1+Credit Use)		28,668,016	2.097	3.977	1.433	0	0	10.37
Ln(1+Credit Use)	*Credit Use* >0	6,372,984	9.434	1.389	0.160	0	0	9.314
Panel B: Owner characteristics								
Owner Age	Years	27,852,659	34.3	10.3	1.480	24	31	51
Male	(0/1)	27,856,705	0.528	0.499	−0.112	0	1	1
Married	(0/1)	26,819,021	0.679	0.467	−0.769	0	1	1
Migration	(0/1)	12,252,990	0.448	0.497	0.208	0	0	1
Income	Yuan	28,515,630	6,446	1,681	0.717	4,477	6,261	8,580
Property	(0/1)	27,388,906	0.541	0.402	−0.109	0.040	0.65	0.999
Panel C: Firm characteristics								
Credit Score (CS)		28,668,016	533.6	533.6	0.05	490.6	533.2	577.1
Firm Age	Months	28,665,507	44.2	28.2	0.75	12	38	85
$[Firm\ Age]^2$		28,665,507	2751	3223	1.84	144	1444	7225
Log Sales		28,668,016	9.2	2.3	−1.71	6.9	9.4	11.5
$[Log\ Sales]^2$		28,668,016	89.5	34.8	0.02	47.7	89.1	132.4
Loan Rate	Annual	28,668,016	16.3	0.5	−1.42	15.7	16.4	16.8
$[Loan\ Rate]^2$		28,668,016	267.6	26.0	−0.79	245.9	269.5	283.4
Panel D: Firm performance								
Sales Growth	Monthly	25,855,839	0.88	20.20	389.16	−0.84	−0.14	1.72
Transaction Growth	Monthly	25,931,087	0.92	26.78	325.64	−0.84	−0.16	1.62
Product Growth	Monthly	25,931,087	0.47	146.08	677.01	−0.31	0	0.26
Panel E: Local credit conditions								
PCGDP	Thousand RMB	279	46.85	27.75	1.85	21,34	39.66	87.21
Loan/GDP		262	9.03	14.11	3.74	0.82	4.24	22.39
SOE Empl. Share		314	31.89	14.47	1.23	15.62	30.12	49.15
AvDistance	Kilometers	8,013,856	12.05	28.55	2.62	0.50	1.45	50.89

Table 4: Credit Demand of China's E-Commerce Entrepreneurs

We draw a random sample of 4,014,000 observations from Sample B (= 15%) and estimate three models for (log) online credit use by Chinese e-commerce entrepreneurs: OLS results are reported in Columns (1)-(2), a Heckman selection model in Columns (3)-(4), and a model with city fixed effects in Columns (5)-(6). The dependent variable is defined as the (natural) log of 1 plus the average (time weighted) *Credit Use* in any given month for the period September 2014 to July 2016. At city level we capture credit conditions for traditional bank credit by the loan to GDP ratio (*Loan/GDP*), by employment in state-owned enterprises relative to total employment (*SOE Employment Share*), and the per capita GDP (*PCGDP*). At the more local level we identify the 20 geographically closest bank branches to every firm location (using GPS coordinates) and calculate average distance (*AvDistance*) between any firm and these bank branches in kilometers as a proxy for search costs and information frictions related to traditional banking. The firm-owner characteristics are age (*Owner Age*), dummies for gender (*Male*), marital status (*Married*), migrant status for entrepreneurs born in another province (*Migrant*), and estimates of owner income (*Income*) and property value (*Property*). Controls for firm characteristics are the *Credit Score* and the level (and the squared terms) of *Firm Age*, *Log Sales*, and the annualized *Loan Rate* for online credit. We use ***, **, and * to denote statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:		$10^3 \times \ln(1 + Credit\ Use)$					
		OLS		Heckman		City Fixed Effects	
		Coefficient	T-value	Coefficient	T-value	Coefficient	T-value
		(1)	(2)	(3)	(4)	(5)	(6)
Local credit conditions							
Loan/GDP		−7.360***	−35.18	−8.693***	−41.44		
SOE Empl. Share		5.739***	18.23	1.988***	6.32		
PCGDP	($\times 10^{-3}$)	1.816***	14.40	2.839***	22.47		
AvDistance		8.328***	32.01	5.540***	21.26	3.591***	52.44
Owner characteristics							
Owner Age	Years	−22.320***	−99.86	−24.910***	−111.01	−12.540***	−36.01
Male	(0/1)	5.263	1.23	2.717	0.64	13.700***	110.70
Married	(0/1)	0.617	0.13	0.366	0.08	78.381***	10.57
Migrant	(0/1)	0.253	0.04	0.102	0.02	67.251***	5.83
Income (estimated)		0.097***	65.02	0.107***	71.52	0.136***	58.30
Property (estimated)	(0/1)	2.873	0.51	1.529	0.27	40.070***	44.93
Firm characteristics							
Credit Score		−1.053***	−44.02	−1.304***	−51.55	−0.357***	−3.03
Firm Age	Months	29.190***	111.70	20.700***	66.13	33.610***	81.29
$[Firm\ Age]^2$	Months	−0.225***	−99.85	−0.149***	−58.72	−0.306***	−87.34
Log Sales		−12.210***	−4.19	−6.303***	−1.91	−216.011***	−14.58
$[Log\ Sales)]^2$		−0.101	−0.53	−0.139	−0.61	20.303***	25.71
Loan Rate		−1.476	−0.08	−0.528	−0.03	−597.501***	−19.71
$[Loan\ Rate]^2$		8.091***	12.17	9.183***	13.01	50.160***	48.54
Inverse Mill's ratio				1.887	0.14		
Firm typ dummies		Yes		Yes		Yes	
City Fixed Effects		No		No		Yes	
Observations		4,014,000		4,014,000		4,014,000	

Table 5: Online Credit Use by Credit Scarcity and Credit Score

We report in Panel A the percentage of Taobao firms using online credit and in Panel B average (log) credit use. Firms are sorted first by online credit use intensity in the city-level prefecture of their location. The city-level online credit use intensity is equal to the city fixed effect for online credit use identified in Table 4, Columns (5)-(6). A second sort groups firms into eight bins according to the credit score of the firm, ranging from Low in Column (1) to High in Column (8).

Panel A: Share of Eligible Firms Using Online Credit (in percent)

City-level online credit use intensity	Firm credit score rank								(1)-(8)
	(1) Low	(2)	(3)	(4)	(5)	(6)	(7)	(8) High	
(A) High	43.91	34.46	32.07	29.80	27.37	24.47	21.25	17.29	26.62
(B)	42.89	33.52	31.46	30.47	28.94	25.94	23.32	19.15	23.74
(C)	40.14	32.42	30.02	28.27	26.29	23.80	20.88	17.80	22.33
(D)	40.03	30.90	28.02	25.57	24.25	21.89	18.62	15.57	24.46
(E) Low	36.22	27.54	25.59	23.88	22.08	20.14	17.78	15.23	20.99
(A)-(E)	7.69	6.92	6.48	5.92	5.29	4.33	3.47	2.06	

Panel B: Average Log Credit Use (in RMB)

City-level online credit use intensity	Firm credit score rank								(1)-(8)
	(1) Low	(2)	(3)	(4)	(5)	(6)	(7)	(8) High	
(A) High	4.021	3.204	3.02	2.838	2.635	2.377	2.079	1.731	2.290
(B)	3.854	3.064	2.898	2.831	2.697	2.429	2.205	1.860	1.994
(C)	3.618	2.958	2.755	2.618	2.444	2.236	1.966	1.716	1.902
(D)	3.605	2.814	2.572	2.376	2.277	2.070	1.764	1.515	2.090
(E) Low	3.313	2.553	2.400	2.263	2.119	1.956	1.746	1.542	1.771
(A)-(E)	0.708	0.651	0.62	0.575	0.516	0.421	0.333	0.189	

Table 6: Credit Demand by Firm Credit Score

We estimate a linear OLS regression for (log) online credit use in four subsamples of Sample B sorted by firm credit scores. Columns (1)-(4) repeat the regression in Table 2 for the firm with a credit score in the quantile range $[0\%, 25\%)$, $[25\%, 50\%)$, $[50\%, 75\%]$, and $[75\%, 100\%]$, respectively. The dependent variable is defined as the (natural) log of 1 plus the average *Credit Use* in any given month for the period September 2014 to July 2016. At the city level we capture credit conditions for traditional bank credit by the loan to GDP ratio (*Loan/GDP*), by employment in state-owned enterprises relative to total employment (*SOE Employment Share*), and per capita GDP (*PCGDP*). At the more local level we identify the 20 geographically closest bank branches to every firm location (using GPS coordinates) and calculate the average distance (*AvDistance*) between the firm and these bank branches in kilometers. The firm owner characteristics are age (*Owner Age*), dummies for gender (*Male*), marital status (*Married*), migrant status for entrepreneurs born in another province (*Migrant*), and estimates of owner income (*Income*) and property value (*Property*). Controls for firm characteristics are *Credit Score* and the level (and the squared terms) of *Firm Age*, *Log Sales*, and the annualized *Loan Rate* for online credit. We use ***, **, and * to denote statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:		$10^3 \times \ln(1 + Credit\ Use)$			
Sampling: Firms with $CS \in$		OLS $[0\%, 25\%)$ (1)	OLS $[25\%, 75\%)$ (2)	OLS $[50\%, 75\%)$ (3)	OLS $[75\%, 100\%]$ (4)
Local credit conditions					
Loan/GDP		-9.374^{***}	-11.961^{***}	-7.042^{***}	-3.906^{***}
SOE Empl. Share		2.047^{***}	5.654^{***}	8.655^{***}	9.206^{***}
PCGDP	$(\times 10^{-3})$	6.005^{***}	4.358^{***}	1.673^{***}	1.053^{***}
AvDistance		1.630^{***}	3.445^{***}	3.935^{***}	4.785^{***}
Owner characteristics					
Owner Age	years	-38.310^{***}	-25.290^{***}	-20.200^{***}	-18.680^{***}
Male	(0/1)	2.647	5.044	5.125	1.077
Married	(0/1)	0.709	0.639	0.678	0.295
Migrant	(0/1)	-0.041	0.082	0.096	-0.001
Income (estimated)		0.275^{***}	0.135^{***}	0.117^{***}	0.102^{***}
Property (estimated)	(0/1)	1.861	3.619	3.912	1.145
Firm characteristics					
Credit Score		0.372^{***}	0.529^{***}	-1.276^{***}	0.259^{**}
Firm Age	months	40.550^{***}	38.310^{***}	32.410^{***}	4.416^{***}
$[Firm\ Age]^2$	months	-0.243^{***}	-0.224^{***}	-0.168^{***}	0.007^{***}
Log Sales		-8.552^{***}	-10.870^{***}	-8.125^{***}	-0.829^{***}
$[Log\ Sales)]^2$		-5.888^{***}	5.197^{***}	9.107^{***}	10.790^{***}
Loan Rate		-1.248	-5.338	-3.994	-1.678
$[Loan\ Rate]^2$		6.062^{***}	2.419^{***}	2.883^{***}	-0.221
Firm typ dummies		Yes	Yes	Yes	Yes
Observations		3,934,000	3,934,000	3,953,000	3,941,000

Table 7: Online Credit and Short-Term Firm Growth

We report in Panel A reduced form regressions and in Panel B two-stage least square regressions (2SLS) based on a fuzzy regression discontinuity design (FRDD). The instrument (IV) is based on the (fuzzy) credit approval discontinuity at the credit score (CS) of 480 points. Treatment is defined as the approval of a credit line to firms with a credit score in the range of 460-500 that did not previously have credit access. We define the binary instrument as $IV(CR_{i,t}) = 1$ if $CS_{i,t} \in [480, 500)$ and $IV(CS_{i,t}) = 0$ if $CS_{i,t} \in [460, 480)$. In Panel A, Column (1), we regress (binary) *Credit Approval* on the (binary) instrument. Columns (2)-(4) report reduced form estimates for various measures of firm performance such as short-term *Sales Growth* (from the end of month t to $t+1$) in Column (2), *Transaction Growth* in Column (3), and *Product (Variety) Growth* in Column (4). In Panel B, Columns (6)-(8), we estimate the linear relationship between the same firm performance measures and (instrumented) log credit use [ln(1+*Credit Use*)]. We include third-order polynomial terms of a firm's credit score (CS) as controls and use ***, **, and * to denote statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Reduced Form Regression

Dependent Variables:		*Credit Approval* (0/1) (1)	*Sales Growth* (2)	*Transaction Growth* (3)	*Product Growth* (4)
IV	(0/1)	0.289***	0.052***	0.045***	0.011***
		(63.800)	(4.087)	(3.833)	(2.737)
CS	($\times 10^{-3}$)	5.678***	1.607	1.661*	0.622*
CS^2	($\times 10^{-6}$)	−17.702**	−53.701***	−40.801**	−4.920
CS^3	($\times 10^{-6}$)	−2.770***	0.792	0.564	0.381
Observations		217,800	217,800	217,800	217,800

Panel B: 2SLS Regressions

Dependent Variable:		*ln*(1 + *Credit Use)* (First Stage) (5)	*Sales Growth* (6)	*Transaction Growth* (7)	*Product Growth* (8)
IV		0.159***			
		(4.792)			
ln(1 + *Credit Use)*			0.292***	0.204**	0.031**
(instrumented)			(2.951)	(2.336)	(1.965)
CS	($\times 10^{-3}$)	−10.440***	4.218***	4.609***	1.447***
CS^2	($\times 10^{-6}$)	−0.208***	−44.100*	18.011	5.820
CS^3	($\times 10^{-6}$)	8.060	0.256	−2.990	−0.648
Observations		217,800	217,800	217,800	217,800