Big Data's Role in Expanding Access to Financial Services in China

By: Nir Kshetri

Abstract:

General consumer and business finance companies have had limited success in serving the needs
of economically active low-income families and micro-enterprises cost-effectively and
sustainably in emerging economies such as China. Recent advances in computing and
telecommunications technology are dramatically transforming this landscape by changing the
way the financial industry operates. A key mechanism underlying this transformation concerns
the use of big data in assessing, evaluating and refining the creditworthiness of potential
borrowers and reducing the transaction costs. While China's internet-only banking industry is
currently small and some activities of players in this industry are akin to those in the shadow
banking, this industry has potential to cause a major disruption in the Chinese financial market.
A main objective of this paper is to examine the role of big data in facilitating the access to
financial products for economically active low-income families and micro-enterprises in China.
A second objective is to investigate how formal and informal institutions facilitate and constrain
the use of big data in the Chinese financial industry and market. The paper also investigates how
various inherent characteristics of big data – volume, velocity, variety, variability and
complexity – are related to the assessment of the creditworthiness of low-income families and
micro-enterprises. Case studies of big data deployment in the Chinese financial industry and
market are discussed. The paper also looks at various categories of personal financial and non-
financial information that are being used as proxy measures for a potential borrower's identity,
ability to repay and willingness to repay. Various business models involving the sources of data
(internal vs. external to the big data organization) and providers of credits (big data organization
vs. external partners or clients of the big data organization) are investigated. The analysis of the
paper indicates that the main reason why low-income families and micro-enterprises in China
and other emerging economies lack access to financial services is not because they lack
creditworthiness but merely because banks and financial institutions lack data, information and
capabilities to access the creditworthiness of and effectively provide financial services to this
financial disadvantaged group.

Article:

1. Introduction

Compared to industrialized countries, developing countries such as China exhibit lower penetration of financial services (Honohan & King, 2009). The problem is more acute and difficult for low-income families and micro-enterprises in emerging economies than for high-income families and large enterprises (Clarke, George, Cull, & Martinez Peria, 2001). General consumer and business finance companies and microcredit organizations have had limited success in serving the needs of these groups cost-effectively and sustainably.

The Chinese financial market deserves special attention. Lending in the country is disproportionately oriented toward powerful economic and political interests such as state-controlled companies (Kshetri, 2011). Small and midsize enterprises (SMEs) account for 70% of GDP but have access to only 20% of financial resources (Klein & Cukier, 2009). It was reported that 89% of SMEs in China face difficulty in satisfying banks' requirements in order to get loans (Jing, 2014). Small borrowers often tend to lack sufficient collateral, which is required by most traditional Chinese banks (Wildau, 2015).

Prior researchers have identified two main problems that contribute to the low penetration of financial services among low-income families and micro-enterprises in emerging economies such as China. First, traditional banks are unwilling and reluctant to serve the small-scale borrowers such as poor people and small businesses due to high transaction costs and inefficient processes associated with making small loans to these borrowers (Adams & Nehman, 1979; Rogaly, 1996). The second reason why poor people and small businesses face barriers to access financial products concerns informational opacity (Stiglitz & Weiss, 1981). Part of the problem also lies in the fact that most developing economies are characterized by the lack, or poor performance of credit rating agencies to provide information about the creditworthiness of SMEs. A national credit bureau would collect and distribute reliable credit information and hence increase transparency and minimize banks' lending risks. This situation puts SMEs in a disadvantaged position in the credit market. This is because SMEs tend to be more informationally opaque than large corporations because the former often lack certified audited financial statements and thus it is difficult for banks to assess or monitor the financial conditions (Kshetri, 2014).

Prior researchers have found that different lending behaviors of different groups of banks in terms of the propensity to lend to poor people and small businesses can be explained in terms of the access to information. For instance, Beck, Thorsten, Demirguc-Kunt, and Maksimovic (2004) found that domestic banks had higher degree of willingness to lend to "opaque" borrowers due to the fact that they have more information about such borrowers and better enforcement mechanisms than foreign banks.

How accessibility and affordability of finance can be improved is a pressing policy and theoretical issue that adjoins larger concerns related to poverty alleviation. Recent advances in computing and telecommunications technology are dramatically transforming the financial landscape from the perspective of economically active low-income families and micro-enterprises by changing the way the financial industry operates. Experts say that this problem can be largely eliminated by creating better risk models using increased computing power and new sources of data and information (Baer, Tobias, Goland, & Schiff, 2013). A key mechanism underlying this transformation concerns the use of big data (hereinafter: BD) in assessing, evaluating and refining the creditworthiness of potential borrowers and reducing transaction costs. Some possible data sources include social media and mobile-phone usage patterns and utility-bill payment history (Baer et al., 2013).

There have been some signs of success on this front. BD is evolving as a transforming force that is likely to shape the Chinese banking sector. Chinese internet companies have launched a broad range of financial products and services. The business models of these companies are centered around the utilization of BD. In February 2015, China's largest online consumer lending marketplace, China Rapid Finance announced that it extended pre-approved loan offers of 500 RMB (about US $80) to 50 million consumers, which included pre-screened users of QQ, the online messaging software developed by Tencent. The offers were made based on an analysis of social and financial information (online and offline) in order to predict default rates, limit fraud and estimate borrowers' responses to the offer (businesswire.com, 2015). China Rapid Finance has estimated that 500 million Chinese consumers are potentially suitable borrowers. The company's goal is to reach them using a mobile-based platform to automatically score creditworthiness based on data from diverse sources (Shu, 2015).

China's traditional banks have also recognized that high quality data about customers is a key to succeed in the financial market. These banks are thus taking measures to transform themselves into BD companies. For instance, as of the early 2012, the Chinese financial industry was estimated to have more than 100 terabytes (TB) of structured and unstructured data (IDC, 2012). As of March 2014, Industrial and Commercial Bank of China (ICBC), the country's largest lender, was reported to have over 4.9 petabytes (PB) of data. Likewise, the Agricultural Bank of China (ABC) was estimated to generate 100 TB of structured data and 1 PB of unstructured data in 2014 (ABC, 2014). Similarly, in 2014, the Bank of Communications (BOCOM) reportedly handled about 600 gigabytes (GB) of data daily and had a storage capacity of more than 70 TB (BOCOM, 2014).

In light of the above observations, a main objective of the present paper is to examine the role of BD in facilitating the access to financial products for economically active low-income families and micro-enterprises in China. A second objective is to investigate how formal and informal institutions facilitate and constrain the use of BD in expanding the access of financial services in China.

The paper is structured as follows. We proceed by first discussing the method employed in this study. Next, we provide a review of the relevant literature. Then, we discuss BD's role in increasing access to finance of low-income families and micro-enterprises in China. The section following this provides discussion of the cases. Next, institutional factors affecting BD

deployment in the Chinese financial industry and market are discussed. The final section provides implications and concluding comments.

2. Method

The approach of this study can be described as theory building from multiple case studies, which is becoming increasingly popular in social science (Eisenhardt & Graebner, 2007). A potentially valuable research design to test the conceptual framework via multiple case studies would be to sample organizations that have been identified as engaging in increasing access to financial services for low-income families and micro-enterprises in emerging economies. In a multiple case study design, the choice of cases needs to be made on a substantive rather than statistical basis in order to adequately represent a target population (Greene & David, 1984). The cases selected in this study thus include diverse types of BD firms.

2.1. Data sources

This study mainly relies on archival data which is among a variety of recognized data sources for case studies (Eisenhardt & Graebner, 2007). As suggested by prior researchers (Golder, 2000; Mason, McKenney, & Copeland, 1997), we also analyzed the sources of evidence as well as the evidence by using the criteria developed by Gottschalk's (1969) such as time elapsed between events and reporting, openness to corrections, range of knowledge and expertise of the person reporting the events, and corroboration from multiple sources.

The paper has articulated the underlying theoretical arguments that provide the logical link between the constructs. As suggested by prior researchers (Eisenhardt & Graebner, 2007; Whetten, 1989), the arguments are based on the cases or from other detached logical reasoning and knowledge (e.g., cases that are not explicitly discussed in the next section).

3. Literature review

We structure the literature review around three key aspects of this study: (a) barriers to the access to financial services faced by low-income families and micro-enterprises in emerging economies; (b) the transaction cost economics approach; and (c) informational opacity, moral hazard and adverse selection problems.

3.1. Barriers and challenges related to accessing financial services faced by consumers and entrepreneurial firms

As noted earlier, conventional financial institutions and microcredit organizations have had a limited success in serving the needs of economically active low-income families and micro-enterprises. One estimate suggested that as of 2010, 2.5 billion of the world's adults did not use formal or semiformal financial services (e.g., MFI) (Chaia et al., 2010).

A study conducted by the World Bank indicated that although China's banking systems were among the world's biggest, the country's private entrepreneurial firms faced higher degrees of financial constraints than those in most other countries (Batra et al., 2003). The Chinese

government's surveys of the private sector, which were conducted before 2002, indicated that entrepreneurial firms consistently ranked access to finance as the biggest obstacle facing their survival and growth (Huang, 2005).

Until 1998, the four state-owned commercial banks—the Bank of China (BOC), the People's Construction Bank of China (PCBC), the ABC and the ICBC were expected to lend only to state-owned enterprises (SOEs). Smaller credit cooperatives were the primary lending institutions for private enterprises (Lin, 2011). Park and Sehrt (2001) found that Chinese financial institutions' lending behaviors were motivated primarily by political considerations instead of economic fundamentals.

Some key developments in these areas have failed to improve this condition. For instance, a rapid increase in the degree of foreign participation has been among the key transformations undergoing the banking sector in developing economies since the mid-1990s (Cull & Martinez Peria, 2007). This trend is associated with and facilitated by a growing trend towards globalization and financial integration. For instance, a study of banking sector assets in 104 developing countries indicated that during 1995–2002, the average share held by foreign banks increased from 18% to 33% (Micco, Panizza, & Yañez, 2007). However, large foreign banks have often exhibited a tendency to abstain from lending to SMEs (Clarke et al., 2001).

In China's case, Lin (2011) found that liberalization of the banking sector, which allowed the entry of foreign bank into the country, alleviated financial constraints facing entrepreneurial firms. The effect was especially pronounced among firms that were less connected to the government. In this way, foreign banks, to some extent, helped reduce the inefficiency in resource allocation associated with state-owned banks' discrimination against private firms (Lin, 2011). The effect; however, remains comparatively small. Foreign lenders controlled only 1.7% of total banking assets in 2013 (Wildau, 2015).

China implemented major initiatives following the 2008 global financial crisis (GFC). In November 2008, the Chinese government announced a fiscal stimulus package of US $586 billion. China also announced a substantial monetary stimulus, which included eliminating lending quotas and reducing interest rates at a four- year low (China Country Report, 2009). These measures stimulated bank lending and led to an increase in prices of shares and commodities. In 2009, Chinese banks lent US $1.4 trillion, which was twice the 2008 level, and half the GDP (Xinhua, 2010).

While the above progress is impressive, Chinese SMEs face difficulties in accessing financial services. Elliott et al. (2015, p. 7) note: "… despite progress that has undoubtedly been made, even optimistic analysts generally agree that SMEs [in China] remain at a considerable disadvantage with banks, over and above those confronted in other countries".

3.2. The transaction cost economics approach

Big state banks have dominated the Chinese financial market with huge networks of branches across the country. For instance, in 2015, ABC had about 24,000 branches, ICBC had about

18,000, China Construction Bank (CCB) had about 13,000 and BOC had about 11,000 (Hongyuran et al., 2015). These banks; however, often find SMEs as unattractive borrowers.

The transaction cost economics approach (Williamson, 1989) can provide insights into barriers faced by SMEs and low income population in accessing financial products and services from conventional financial institutions. A reason why traditional banks are unwilling to serve the small-scale borrowers is that these borrowers are characterized by high transaction costs and inefficient processes.

In an analysis of farm level information from Bangladesh, Brazil and Colombia, Adams and Nehman (1979) found that small borrowers' borrowing costs on formal loans, as defined by the sum of the nominal interest payments, borrower loan transaction costs and changes in the purchasing power of money, were substantially higher than those of large borrowers. Prior researchers have argued for creative and innovative designs in financial services in order to reduce the transaction costs making small loans to poor people and small businesses (Rogaly, 1996).

The Internet has drastically reduced transaction costs associated with financial and banking activities. For example, the average cost of a banking transaction is estimated to be US $1.27 in a branch and US $0.27 in an ATM, whereas it is US $0.01 on the Internet (UNCTAD, 2000). A study indicated that, for a transaction involving US $23, branchless banks cost 38% less than commercial banks and 54% less than informal money transfer channels (McKay & Pickens, 2010). For instance, the average mobile transaction conducted via the mobile payment system, M-Pesa was reported to be about a hundredth of the average check transaction and half of the average ATM transaction (Jack & Suri, 2010).

3.3. Informational opacity, moral hazard and adverse selection problems

The barriers faced by small firms to access financial products from conventional financial institutions are the often result of informational opacity, which may lead to moral hazard and adverse selection problems (Stiglitz & Weiss, 1981). In China, only 20% of the adult population has a credit score. They often get credits from large SOEs through the People's Bank of China (Lohr, 2015).

Prior researchers have provided evidence for an important role of information in facilitating the development of the financial market and access to financial products of a broader range of market participants (Beck et al., 2004). Jappelli and Marco (2002) suggested that the degree of information sharing between intermediaries is positively related to financial development.

4. Big data's role in increasing access to finance of low-income families and micro-enterprises in China

4.1. Data to assess creditworthiness of potential borrowers and reduce transaction costs

A wide array of digital activities generate a huge amount of structured and unstructured data. Table 1 presents some indicators that give an idea about the amount of personal data in China. It

is worth noting that China is the biggest market in terms of mobile phone users, Internet users, and social media users. Table 2 presents some indicators related to the size of consumer data held by China's digital conglomerates. It is clear that these companies have a massive amount of data on online consumption.

Table 1. Some indicators related to the market sizes of computer-based information technologies in China.

Indicator	Statistics	Source
Mobile phone users (June 2015)	675 million unique e users (1.3 billion SIM subscriptions).	The Ministry of Industry and Information Technology (MIIT) (Chinadaily.com), CNNIC
	1.29 billion (94.5% of the population)	
Active internet users (August 2015)	668 million	CNNIC (Milward, 2015)
Active social media users (August 2015)	659 million	CNNIC
Mobile transactions (2013)	No. of transactions 1.668 billion	The China Banking Regulatory Commission (CBRC, 2014)
	Mobile payments 9.64 trillion yuan (US $1.6 trillion)	
	No of. mobile banking customers: 458 million (annual growth 55.5%).	
Consumers' ecommerce spending (first half of 2015)	US $253 billion (10% of total retail sales)	CNNIC
Machine to machine (M2M) connections	Mid-2014: >50 million (over a quarter of the global total)	GSMA (GSMA, 2015)
	Mid-2015: 74 million (the world's largest M2M market)	

Table 2. Some indicators related to the size of consumer data held by China's digital conglomerates.

Company	User-related data and statistics	Financial-related data and statistics
Alibaba	Early 2015: over 300 million registered users and 37 million small businesses on Alibaba Group marketplaces including Taobao and Tmall.com (alibabagroup.com, 2015).	Fourth quarter of 2014: revenue $4.2 billion (Alibaba Group, 2015)
	Early 2015: Taobao had over 500 million registered accounts and over 7 million merchants, which sold 4,800 items per minute (forbes.com, 2015).	
Tencent	Early 2015: QQ had over 800 million users.	Total revenues: US $7.5 billion (first half of 2015) (Tencent.com, 2015)
	Mid-2015: WeChat had 550 million active users.	
JD.com	Mid 2015: 100 million active customers (Lohr, 2015).	Annual revenue: $20 billion

4.1.1. Data generated by the internet of things

In the future, the data generated by the internet of things (IoT), the network of physical objects or "things" (e.g., machines, devices and appliances, and animals or people) embedded with electronics, software and sensors, which are provided with unique identifiers and possess the ability to transfer data with minimal human interventions, may also be used to assess creditworthiness of individuals and organization. In this regard, China is already the world's largest market for machine to machine (M2M) connections, a subset of the IoT, which use wireless networks to connect devices to each other and with the Internet.

The IoT is referred as the "third industrial revolution" (Modak, 2015). According to PWC's 6th annual digital IQ survey, financial services is among the top 10 industries in terms of the investment in sensors as potential IoT innovations (Modak, 2015). According the "Internet of Things in Banking" study conducted for Verizon by American Banker and Source Media Research in August 2014, 13% of banks had implemented a M2M solution (McGehee, 2015). According to one study, 54% of the financial industry's top performers plan to increase investments in sensors in 2015 (Conlan, 2015). For emerging economies such as China, the IoT's potential to deploy in the financial industry and market stems from the fact that the costs of sensors have reduced dramatically. For instance, RFID tags cost just a few cents (McLellan, 2013). For instance, EPCglobal, whose goal is to achieve worldwide adoption and standardization of Electronic Product Code (EPC) technology, offered its silicon-based tags, which can store a unique serial number, for 5 cents per tag (RFIDJournal).

Financial institutions can harness the power of IoT in several ways. For instance, knowing that a clients' washing machine is likely to break down in the next few weeks would allow retail banks to offer a good credit deal earlier than competitors. Predicting the number of appliances that need

maintenance would allow extending loans to small businesses with a low risk (Schutzer, 2014). With the IoT, banks and credit unions will also be in better a position to offer personalized financial and nonfinancial advice to consumers. When a customer, whose car financing has been approved visits a dealer, a bank can use the IoT to alert the consumer regarding the approved amount or deliver customized loan proposals in a timely manner (Modak, 2015).

Financial institutions can also facilitate consumers' access to insurers and other related service providers. Homeowners can voluntarily provide data related to household management. Smart home appliances such as water leakage detector and air purifiers are likely to lower property claims and reduce the loss ratio. Likewise, sensors to predict natural disasters such as tsunamis and earthquakes are likely reduce the causality and property loss and lead lower insurance claims (Arunwadkar, 2015). By using sensors, real-time feedback and predictive analysis of behavioral data, it is possible to shift the property and casualty (P&C) insurance business from a "reimbursement" model to a "prevention and loss control" model. Additional data provided by sensors and devices can be supplemented with humans advice on safety and prevention, which is likely to provide a better and more accurate understanding of exposure, hazards and risk. It is thus possible to offer loss control on an individualized scale (Schutzer, 2014). Insurers can reward behaviors such as locking doors when leaving homes, or turning off stoves and ovens when not in use. In the non-IoT environment, there is no measurable or reliable way for home insurers to know this information. The Internet of Things and Smart Homes allow this data to be collected, managed and shared (Modak, 2015).

In the financial industry the IoT will reduce the cost of monitoring a loan in asset-based lending, which is tied to inventory, accounts receivable, machinery and equipment. Note that asset based loans tend to cost more than traditional loans and sometimes due primarily to additional audit and due diligence fees. With an agreement between the lender and the borrower, the IoT can monitor and track various characteristics of the equipment and/or the inventory. This can reduce the cost of monitoring for the lender and reduce the overall risk. The IoT can provide advance warnings on cash flow issues that may lead to default. The IoT thus provides the ability to perform proactive collection and enforce outstanding payments. Borrowers can be offered lower rates if they opt-in to be tracked by IoT sensors on equipment and inventory. Note that a similar strategy is already used in some personal auto loans, in which an IoT can help in the repossession if the borrower defaults. In auto-insurance, the IoT can provide valuable driving performance information, which can be used by insurance companies to provide discounts to drivers (Hartani, 2015).

4.2. Selected cases of BD deployment in the Chinese financial industry and market

4.2.1. Alibaba

Alibaba uses BD to improve risk management and control. Alibaba has developed its own credit ratings and risk control models based on information on payment and e-commerce transaction. It mainly utilizes its huge online ecosystem, which, as of the early 2015, consisted of over 300 million registered users and 37 million small businesses on Alibaba Group marketplaces including Taobao and Tmall.com (alibabagroup.com, 2015).

In 2007, Alibaba launched AliLoan, which was backed by the CCB. The CCB had a lot of money to lend and was looking for attractive borrowers but was cautious of lending to small businesses that lacked credit histories (Rabinovitch, 2013). AliLoan focused on small companies. Alibaba provided transaction data from its ecommerce site to the CCB so that the later could make better-informed lending decisions.

The relationship terminated in 2011 when Alibaba reportedly asked the CCB to pay more for its credit information. Alibaba then used its own funds to lend via its AliFinance website. By mid-2012 AliFinance extended Rmb 28 billion in loans to more than 130,000 small businesses (Rabinovitch, 2013). During the three-year period following its creation, AliFinance issued over Rmb 100 billion ($16 billion) of credit. As of February 2014, the company had made loans worth over 170 billion yuan to more than 700,000 SMEs (Li et al., 2014). Every loan is estimated to cost 0.3 yuan or roughly one-thousandth of what a traditional loan by a bank would cost (Li, 2014). Its nonperforming loan ratio was reported to be less than 1% (Cui, 2013).

In June 2013 Alibaba launched new fund management service called Yu'e Bao ("extra treasure" in Chinese) to compete with banks' deposit business. Yu'E Bao allows e-commerce customers to deposit leftover cash into a high-interest fund. A few months after it was launched, the service attracted over US$1 billion in investments (economist.com, 2013). Yu'E Bao is an attractive alternative to traditional bank accounts, which yielded annual interest rates of about 4.5%, which was significantly bigger than the 0.35% rate on current deposits in banks but enjoy the same liquidity. That is, funds can be withdrawn at any time (Rabinovitch, 2013). Yu'e Bao attracted 574 billion yuan of funds by June 2014 (Chen, 2014).

In April 2014, Alibaba started the platform Zhao Cai Bao, which lets small businesses and individuals to borrow from investors directly. It created a 14 billion yuan (US $2.3 billion) marketplace as of September 2014. A borrower can get loans from up to 200 investors after a financial institution guarantees the loan and makes sure the money will be paid back. Zhao Cai Bao worked with over 40 financial institutions to help guarantee the credit (Chen, 2014).

In July 2014, Alibaba launched the Open Data Processing Service (ODPS), which allows users to remotely tap into Alibaba servers equipped with algorithms. According to Alibaba, the system had the capability to process 100 million high-definition movies' worth of data in six hours (Li, 2014). The program uses more than 100 computing models to process over 80 billion data entries every day. A vendor's willingness and ability to repay loans are assessed based on information such as the borrower's credit rating and customer reviews. As of 2014, more than 70 people worked on developing models for ODPS in the small-loan business. All decisions related to granting a loan are made by the system without human intervention (Li, 2014).

Sesame credit was launched in January 2015, which provides credit ratings of consumers and small businesses. Sesame mainly utilizes data from Alibaba's huge online ecosystem. It also makes use of BD collected from Alibaba's various partners, as well as the online and offline history of transactions. Sesame provides credit providers with more accurate and data-driven insights, which can help assess potential borrowers' creditworthiness and offer loans and microfinance and other credit-related services.

In June 2015, Ant Financial Services Group, the financial affiliate of Alibaba Group Holding, launched MYbank, which is an Internet bank run entirely on the cloud. Rural users are one of key target groups of MYbank (Asia Unhedged, 2015). The company aspires to provide credits to farmers for buying agricultural machines and tools (Alois, 2015). Ant financial plans to issue loans of up to US $800,000 to small businesses and consumers (Rao, 2015). MYbank's goal is to extend loans to 10 million SMEs in five years (Bloomberg.com, 2015). MYbank's data comes Ant Financial as well as from credit evaluation companies, such as Zhima Credit (Alois, 2015), which is Alibaba's personal credit scoring service launched in January 2015. In determining a user's credit score Zhima uses information such as court reports default on debts, late returns of rented cars and transactions on Alipay online payment service (wantchinatimes, 2015).

Alibaba is exploiting its massive amount of data related to online consumption in the offline setting by accessing the relevant data through Alipay (China.org, 2015). Alipay credits can be used to buy consumer goods with offline retailers. In June 2015, Ant Financial announced a partnership with Wal-Mart stores in which the latter would accept the Alipay mobile payments service. The partnership would start with 25 stores in Shenzen and cover all 410 Wal-Mart stores in China by the end of the 2015 (Wahba, 2015). Alibaba's BD research on customer preferences, behavioral habits and credit ratings is expected to help Walmart to better utilize consumer profile information. This is expected to help the company launch more effective marketing promotions and reduce operational costs (china.org, 2015).

4.2.2. Tencent

Tencent's huge online ecosystem consists of the users of WeChat, which offers a number of features such as celebrity marketing, free calls and texts and QQ, the online messaging software. In the mid-2015, WeChat was reported to have about 550 million active users, which was about twice that of Twitter (Davidson, 2015). As of the early 2015, QQ had over 800 million users (Zhou, 2015).

In the early 2015, Tencent-backed private lender, WeBank was launched. WeBank carries out all services online and it has no plan to have brick-and-mortar outlets. A key aspect of WeBank's financial innovations involves BD. To apply for a loan, an applicant takes her/his picture using the cellphone camera and submits online. WeBank's system matches the picture with data provided by the Ministry of Public Security in order to verify the person's identity. The system gives the applicant a credit rating based on data collected from a number of sources such as online shopping, activities on social network websites games and other online activities. The credit rating is used to decide the amount of credit that the applicant can obtain (Zhou, 2015).

Tencent's key strategy has been to make credits available to regular WeChat users in order to encourage them to shop at selected retailers. By the end of 2014, over 100 million users' bank cards had been linked with Tencent's payment system through WeChat and QQ (Yining, 2015). The firm is planning to launch a point schemes and membership plans, which will give WeChat users a better deal, if they use the WeChat payment system (china.org, 2015).

4.2.3. JD.com (Jingdong mall)

JD.com is a Chinese online retailer. As of the mid 2015, it had 100 million active customers. It is reported to generate yearly revenue of $20 billion (Lohr, 2015). In June 2015, ZestFinance and JD.com announced a joint venture, JD-ZestFinance Gaia, to provide a consumer credit scoring service in China. In early stages, JD-ZestFinance will focus on assessing credit risk and offer installment loans for purchases on JD.com. In the future, the venture aims to offer credit-analysis service to corporate customers in China (Lohr, 2015).

ZestFinance reportedly uses "tens of thousands of data points" to assess potential borrower's ability to pay back loans. Consumers who want to buy items on JD.com can apply for a line of credit. ZestFinance and JD.com utilize data from the applicants' past and present online shopping habits to predict the risk of default. ZestFinance's algorithms also take into account factors such as time of day the person is shopping online (which is used as an indication of whether the applicant has a job), the type of items being purchased, and the history of buying expensive items (Rao, 2015).

4.2.4. China CITIC bank

China CITIC bank has established a database that consists of detailed customer profiles including demographic information (e.g., transaction data, attributes of customer, residence information, etc.) and online behavior information (e.g., transaction and risk preferences, social networking activities, etc.). It uses the information to improve its precision marketing and better target credit-card customers.

The bank's credit approval process in the past used only indicators such as income and employment status, which led to the rejection of a large proportion of qualified customers. The rejection was due to its lack of ability to accurately assess the risk levels. It was reported that for new customers who applied through the Bank's online channels, the approval rate increased from about 25% in the past to about 70–80% (Chng & Zhang, 2014).

4.2.5. China's state-owned banks

Chinese traditional banks have understood how valuable and important customer information is. In the Alibaba-CCB conflict discussed above, the CCB responded by establishing its own online mall to collect transaction data. In June 2012, CCB launched an online mall, buy.ccb.com. The goal was to become a fully-fledged ecommerce site (Rabinovitch, 2013). In the six months since its founding, CCB's ecommerce site registered 10,000 stores and facilitated Rmb 3.5 billion (about US $500 million) in transactions (Rabinovitch, 2013). In the first half of 2015, the site was reported to attract 57,000 vendors and about 12 million shoppers, who mad 30 billion yuan worth of transactions (Hongyuran et al., 2015).

The BOCOM opened an online shopping mall in September 2012. Likewise, in January 2014, the ICBC launched e-commerce website. By June 2014, the service generated 200 billion yuan in sales, which ranked third in terms of revenue, behind Alibaba's tmall.com and JD.com (Hongyuran et al., 2015). In October 2013, BOC launched its e-commerce site. As of September 2015, the website had about 15,000 vendors and 1.2 million users. Likewise, in April 2015, ABC launched the e-commerce site www.e.abchina.com, which targets people in rural China. As of

June 2015, the website had 20,000 vendors, which conducted sales worth 700 million yuan (Hongyuran et al., 2015). These bank's e-commerce businesses generate demands for financial products from consumers as well as businesses (Hongyuran et al., 2015).

These banks are forced to be efficient by shifting businesses online. A caixin.com article quoted a bank executive saying that up to 95% services being offered at branches can be handled online (Hongyuran et al., 2015). As a strategic response to the emergence of web-only banks, in September 2015, the ICBC announced a new "Internet financing center", which was a part of the bank's e-ICBC framework. Other key components of the framework include an online financial information platform, an e-commerce platform at www.mall.icbc.com.cn, and an online banking platform (Hongyuran et al., 2015). The bank Chairman noted that he expected 100 million customers would use ICBC's online business platforms in the near future. He also predicted that ICBC will be a leading player in the online financing market (Hongyuran et al., 2015). Other big banks are also expanding into the Internet-based financing market. For this purpose, they are entering into e-commerce businesses. In 2012, the CCB started e-commerce, online banking and mobile banking platforms. The ABC started an e-commerce business, which mainly focuses on consumers in rural areas. The BOC is also reported to have plans to build an online financial services platform (Hongyuran et al., 2015).

4.2.6. The Chinese government's planned credit code

China's State Council has issued guidelines for issuing a Social Credit System by 2020 (chinacopyrightandmedia.wordpress.com, 2015). By then, every adult will have a credit code in addition to a government-issued identity card. A person's credit code is derived from factors such as financial standing, criminal record and social media behavior (FlorCruz, 2015). BD will be a key enabler in the construction of this system. This is perhaps the most spectacular example of a deployment of BD analytics and technology in governance and one which illustrates a truly powerful and provocative way to manipulate and control the population.

The stated goal of the plan is to contribute to constructing a harmonious socialist society (chinacopyrightandmedia.wordpress.com, 2015). Its broader aims include the promotion of positive economic (e.g., avoiding extravagant consumption) and moral (promotion of socialist core values such as patriotism, respecting the elderly and working hard) behaviors. Individuals with a bad credit are likely to be subject to various sanctions such as financial penalties (e.g., no housing or credit to start a company) and ineligibility for certain jobs (volkskrant.nl, 2015). High-status and influential positions such as government official, news reporter, and CEO of a company may require a certain minimum score.

Some foundation for Social Credit System has already been laid. In June 2015, the website of Credit China (creditchina.gov.cn) was launched, which contained more than 1.13 million pieces of information on credit histories of some Chinese residents and firms such as tax avoidance and failure to follow court rulings (Xinhua, 2015). The website allows anyone to check anyone else's credit score. Credit China's information comes from various government departments such as the State Administration of Taxation, Supreme People's Court, National Development and Reform Commission (NDRC) and the China Securities Regulatory Commission and local governments (http://english.gov.cn/policies/latest_releases/2015/07/01/content_281475138273106).

Publicly questioning or challenge the viewpoint of the Chinese Communist Party (CCP) would lead to a lower credit score. For instance, behaviors such as that lower one's score were reported to include posting political opinions without the government's prior permission and describing a different history than what the government has told would degrade the credit rating. Publishing anything that annoys and embarrasses the government such as the news related to the 2015 Shanghai stock market collapse would also lower the score. It is argued that information included in the rating may also include what books people read. An expert referred it as "Amazon's consumer tracking with an Orwellian political twist" (volkskrant.nl, 2015). Moreover, a person's score is also a function of political opinions held by and behaviors of the person's friends and acquaintances.

The credit system is also likely to benefit from the nascent but rapidly developing credit rating systems of China's private companies. China has given ten private companies permission to launch internet-based credit rating and ranking systems, which include Alibaba's Ant Financial and Tencent's paywall (Huang, 2015).

5. Discussion of the cases

So far, BD companies' share is vanishingly small in China's lending market, which is estimated at US $4 trillion (Bloomberg.com, 2015). For instance, in 2013, credit offered by Alibaba was only 0.1% of the total loans issued by China's banks (Rabinovitch, 2013). Nonetheless, BD companies have started to powerfully influence the Chinese financial market. Chinese banks' interest in acquiring customer data is an encouraging sign that may indicate that lending decisions are likely to be made based on economic fundamentals rather than by political considerations. However, they have witnessed various levels of success and failure. For instance, as noted earlier, ICBC's e-commerce website ranked third in terms of revenue. Nonetheless, many small banks which opened online shopping websites have been forced to close (Hongyuran et al., 2015).

Some activities of players in the Chinese internet-only banking industry are akin to those in the shadow banking. According to Former Federal Reserve Chair Ben Bernanke, shadow banking "comprises a diverse set of institutions and markets that, collectively, carry out traditional banking functions—but do so outside, or in ways only loosely linked to, the traditional system of regulated depository institutions" (Bernanke, 2013). China's BD companies are carrying out an increasing number of traditional banking functions. These players, however, mostly remain outside the regulated sector. For instance, as noted earlier, AliFinance makes loans much bigger than its registered capital would allow in the regulated banking sector (Rabinovitch, 2013).

China's BD companies are likely to attract customers from the existing shadow banking sector. Some estimates suggested that the shadow banking sector accounted for about 40% of new lending in 2012 (Vergis, 2014). According to the Global Shadow Banking Monitoring Report 2014, China's shadow banking sector grew by 38% to reach about $3 trillion (Vergis, 2014). BD companies offer many of the benefits of the shadow banking sector. For instance, a major benefit of shadow banks is that they often make financial services cheaper and more widely available (Elliott et al., 2015). The Yu'E Bao example is illustrative of a BD company's attempt to make

financial services more attractive to consumers that previously lacked access to financial services.

5.1. The relevance of different characteristics of BD

It is first important to explain what BD is and how one can say that a dataset used for accessing creditworthiness and making access to financial products to low income families and micro-enterprises is BD. Gartner has defined BD in terms of three vs: volume, velocity and variety. The software company, SAS, has added two additional dimensions: variability and complexity (sas.com, 2013). The following discussion will examine how the various characteristics or dimensions of BD identified by Gartner and SAS are relevant in the context of facilitating the access to financial products for economically active low-income families and micro-enterprises in emerging economies (Table 3).

Table 3. Relevance of BD dimensions in enhancing access to financial products for economically active low-income families and micro-enterprises in emerging economies.

Characteristic	Explanation	Some examples in the context of assessing creditworthiness of poor people
Volume	Enormous amounts of data created and gathered from a wide range of sources	Alibaba's ODPS processes > 80b data entries/day.
		March 2014: ICBC: 4.9 PB of data. BOCOM: 600 GB daily, storage capacity > 70 TB.
Velocity	Time-sensitive data- collected, stored, processed, analyzed and acted on quickly (speed's importance).	Alibaba: decisions made by the system without human intervention: increases the speed of decisions.
Variety	Data in multiple formats: structured and unstructured.	Structured data (e.g., no./amount of transactions), unstructured data (e.g., WeBank: pictures with the Ministry of Public Security: identity).
		ABC: 100 TB of structured data and 1 PB of unstructured data in 2014.
Variability	Data flows: peaks and troughs	February 7, 2014: daily peak of taxi rides via WeChat: more data flow.
Complexity	Data from multiple sources: linking, matching, cleansing and transforming across systems.	Matching/linking data from various sources:
		WeBank: pictures with the Ministry of Public Security, combining data from other sources.
		Alibaba: various partners, online and offline history of transactions.

		Social credit system: information from various government departments.
		Zhima: information such as court reports default on debts, late returns of rented cars and transactions on Alipay online payment service.

5.1.1. Volume

There has been a colossal increase in the digitization rate in China, which has led to the availability of a large amount of data and information to assess the creditworthiness of individuals and enterprises. For instance, China's internet companies such as Alibaba and Tencent have huge amount of proprietary data related to online purchasing habits and creditworthiness of businesses and consumers (economist.com, 2013).

5.1.2. Velocity

Some data is time-sensitive for which speed is more important than volume. Data needs to be stored, processed and analyzed quickly. The creation of high velocity data has helped increase the access to financial services among the poor. For instance, Alibaba's systems make lending decisions without human intervention, which increases the speed with which such decisions can be made.

5.1.3. Variety

It is important to first define structured and unstructured data. Structured data can be organized in an assigned format that can be used by a database management system such as Oracle and Microsoft SQL. Some examples include histories of mobile payment transaction and the date of a Twitter account creation. Such data can be arranged in a list. Unstructured data, on the other hand, is unformatted and lacks a predefined standard structure (e.g., cannot be organized in terms of rows and columns). Some examples include email messages, social media posts, pictures and video. Structured and unstructured data are being increasingly combined in accessing creditworthiness and making access to financial products to low income families and micro-enterprises. For instance, a technology developed by the Brazilian company, Cignifi can recognize patterns in consumers' phone-calls, text messages and data usage, which are used to predict lifestyle and credit risk profile (bigdata-startups.com, 2013). While various categories of structured data such as amount of transaction are obviously useful and necessary for assessing creditworthiness of potential borrowers, China's BD companies and traditional banks are increasing utilizing unstructured data for this purpose.

5.1.4. Variability

Data flows can vary greatly with periodic peaks and troughs. These are related to social media trends, daily, seasonal and event-triggered peak data loads and other factors. To take an example, on February 7, 2014, there was a daily peak of 2.62 million taxi rides via WeChat, about two

million of which were paid through WeChat payments. Many people were returning home after visiting relatives for the Chinese New Year (Millward, 2014).

5.1.5. Complexity

It would be helpful to first note that the really big difference between variety and complexity concerns multiple data types (variety) versus multiple sources of data (complexity). Matching and linking data from multiple sources such as social media, call detail records (CDRs), open portals, government sources, BD corporations can provide a whole picture of the economic and social conditions of potential borrowers.

5.2. Utilizing different categories of financial and non-financial information

In attempting to explain the potential uses of BD in assessing creditworthiness of poor people and microenterprises, analysts have suggested the importance of paying greater attention to three categories of data that can be potentially used as reliable proxies for creditworthiness: identity (e.g., information that would help reduce frauds), ability to repay (e.g., income, current debt load), and willingness to repay (e.g., past credit experience) (Baer et al., 2013). Table 4 presents the functions of different categories of financial and non-financial information used by lending organizations. Potential sources of data and information about a borrower include telecommunications providers, utility companies, wholesale suppliers, retailers, and government agencies. In addition, financial institutions might have data that were previously overlooked (Baer et al., 2013). Regarding identity, it is worth noting that information and data often come from a third party. Such information generally has a higher degree of trustworthiness than information that comes directly from the potential borrower.

Table 4. Functions of different categories of financial and non-financial information.

	Financial information	**Non-financial information**
Identity (to reduce fraud)		• From government agencies (WeBank from the Ministry of Public Security). • Third party information: high deg. of trustworthiness.
Ability to repay	• Income/current debt. • Prepaid minute purchase patterns: steady/uneven cash flow. • Estimates of future production (collateral).	• Reputation of friends. • Hobbies (e.g., time spent playing video games). • Time of day the person is shopping online.
Willingness to repay	• Past credit experience. • Utility-bill payment history.	• Type of items purchased, (e.g., buying diapers = responsibility, the history of buying expensive items). • Customer reviews (for org.). • Books read.

5.3. Business models used by BD companies

Table 5 presents business models associated with BD use in enhancing access to finance in terms of the dominant sources of data used (internal vs. external) and whether the provider of credits is the BD organization or an external partners or clients of the BD organization. The cases of China CITIC Bank and other traditional banks suggest that banks are increasingly recognizing the importance and relevance of information about potential borrowers.

Table 5. Business models used by BD companies in enhancing access to finance.

Loans/financial services provided by	Dominant sources of data used by the processor of BD	
	Internal	**External**
BD organization (processor of BD and provider of financial services as the same organization	[I]	[II]
	• Chinese government's planned credit code • Alibaba's Mybank • AliFinance • Tencents' WeBank	• China rapid finance • China CITIC Bank
Financial institutions relying on BD organizations' expertise and experience	[III]	[IV]
	• Alibaba's Zhao Cai Bao • AliLoan	• Cignifi

Both BD companies and financial institutions are taking initiatives in order to be in cell [I]. For this purpose, they are invading each other's territories. For instance, banks are diversifying their businesses by investing in e-commerce. E-commerce companies such as Alibaba, on the other hand, have launched financial products and services such as money market funds and consumer loans. Nonetheless traditional banks may find it difficult to achieve the level of BD capability as Alibaba's. Alibaba is at the cutting edge of BD and its activities have been potentially copycat models for a number of Chinese banks. Alibaba's major competitive advantage is in its huge base of clients, extensive transaction records and BD analysis techniques. The company is exploiting these advantages in the lending industry.

5.4. The role of BD in reducing information opacity and transaction costs

Table 6 presents the role of BD in reducing information opacity and transaction costs and compares the approaches of some of the BD firms discussed in the last section. As it is clear from Table 6 diverse sources of data are being used to reduce information opacity. Transaction costs are reduced by digitizing the activities and/or minimizing or completely eliminating physical interaction between the lender and the borrower.

Table 6. The role of BD in reducing information opacity and transaction costs.

	Reducing information opacity	**Reducing transaction costs**
Alibaba	• Proprietary data: online purchasing habits/creditworthiness • ODPS • BD from various platforms/partners, online/offline history of transactions • Scoring system-online/offline data	• No human intervention • My bank: internet only bank
Tencent	• An applicant gets a credit rating based on data collected from a number of sources: used to decide the amount of credit that	• WeBank carries out all services online
JD.com	• Data from the applicants' past and present online shopping habits and factors such as time of day the person is shopping online, the type of items being purchased, and the history of buying expensive items • Zest Finance's algorithms	• All activities are performed online

China's BD companies; however, may not be in a position to attract some of the big and more stable clients, which may require specialized services such as custodial services for safeguarding financial assets, cross-currency settlement and face-to-face advice (Rabinovitch, 2013). Information opacity and transaction costs may not be of concern for these big clients.

6. Institutional factors affecting BD deployment in the Chinese financial industry and market

In this section, we provide a brief overview of some of the mechanisms by which formal and informal institutions may facilitate and constrain the use of BD in expanding the access of financial services in China. First, the Chinese government's policy measures to improve the access to financial services have been instrumental in promoting innovational in the financial market. These policies have been especially influenced by the 2008 GFC. The GFC led to a dramatic fall in China's exports, an overcapacity in manufacturing industries, and an increase in unemployment (China Country Report, 2009). It is estimated that 22 million migrant workers lost jobs following the GFC (Jia, 2009). Investor and philanthropist George Soros noted that while the GFC in the US context was "an internally- generated event leading to the collapse of the financial system," it impacted China through "an external shock to exports" (Mu, 2009). Unsurprisingly measures to reduce external vulnerability and emphasis on a consumer-driven economy have been an important change in China's structure of the developmental model in the post-GFC era (Platt, 2010; Zakaria, 2009). During the World Economic Forum Annual Meeting 2010 in Switzerland, the then Chinese Vice Premier Li Keqiang noted that China would explore a new development model (Kshetri, 2011). He reiterated the government's goals of increasing domestic consumption to drive growth, which would provide the cushioning effect from external shocks (Vidaillet, 2010). Furthermore, China has shown a long-term commitment of a fiscal stimulus, which is expected to support the recovery of the private sector and allow a "self-sustaining growth" (Koo, 2009, p. 18). The Chinese government's political agenda has been to stimulate consumer spending in order to reduce its dependence on industrial exports. Expansion

of consumer credit can play a major role in this endeavor. The government is thus allowing private companies such as Alibaba, Tencents and JD.com to innovate (Lohr, 2015).

It is also interesting to look at the institutional foundation that has supported the growth of BD companies. It is worth noting that the success of China's biggest Internet giants such as Alibaba and Tencent could be attributed to China's Great Firewall. Since foreign Internet companies such as YouTube, Twitter and Google have been blocked in China, Chinese Internet companies have had the space needed to grow (wsj.com, 2015). An article published at Global Times noted that without the firewall "China would become the realm of Google China, Yahoo China and Facebook China" (wsj.com, 2015). A related point is that compliance with the government's censorship policies forms a major component of product development costs for China's Internet start-ups. An upshot of this is that major Internet markets in China are dominated by the monopoly of firms such as Alibaba and Tencent. This means that firms that are endowed with a monopoly position in the market such as Alibaba and Tencent are in a position to attract a large number of customers on their websites and collect a vast array of information about them.

Another way to view this is to think of BD companies such as Alibaba as institutional entrepreneurs and their roles in creating market-oriented institutions. BD companies such as Alibaba are reported to be involved in potentially violating laws. For instance, quoting China Entrepreneur magazine, a FT article noted that, if Alibaba had followed the rules on leverage for small-loan companies, AliFinance's registered capital would not have allowed it to make loans worth more than Rmb 2.4 billion. As of 2013, AliFinance made loans 40 times bigger than this value (Rabinovitch, 2013). Daokui Li, Feng, and Jiang (2006) have identified four approaches institutional entrepreneurs employ to create market-oriented institutions by breaking regulative barriers: open advocacy, private persuasion, making a case for exceptions, and ex ante investment with ex post justification. First, open advocacy works only if the government is— tolerant enough for opinions that may criticize existing policies, regulations or laws‖ and the advocated changes are perceived to be beneficial to the general public (Daokui Li et al., 2006). Second, entrepreneurs may persuade policy makers privately. Third, an entrepreneur may argue that(s) he has a special case that is an exception to the existing laws and regulations. Finally, if a business formed or expanded by breaking existing laws generates jobs, tax revenues and other forms of social benefits, the entrepreneur reports to the government and persuades policy makers to bring changes in existing laws and regulations. China's BD companies have engaged in one or more of the above-mentioned actions. For instance, despite their potential non-compliance with regulations, they have created social benefits by expanding the access to financial services to the broader section of the society. However, some argue that if AliFinance grows to become a big lender regulators may pay more attention to the company's regulatory compliance (Rabinovitch, 2013).

Despite the above observations, companies' that make uses of BD to expand financial services need to tackle some regulatory bottlenecks and impediments. For instance, when Alibaba's online bank, MYbank was launched in June 2015, registration of customers was not possible due to regulatory concerns over the facial-recognition technology used to verify users' identities (chinaeconomicreview, 2015).

One of the striking features of China's government-business relationships is the symbiotic nature that embodies the interdependence and benefits both parties. Internet companies profit from extensive data mining and the government can access the data to improve its surveillance and censorship programs. A deep government-business cooperation already exists in China's censorship and surveillance system, which can be considered to be one of the most effective BD applications. Unlike in the West, Chinese technology firms have openly supported and worked with the government and facilitated the censorship and surveillance programs. Internet firms such as Baidu, Tencent, Sina Weibo and Youku have developed elaborate big data systems for censoring and filtering user generated contents. These companies work with the CCP's propaganda apparatus.

It is a common practice among Internet firms to seek for the government's input before launching a product or service. In this way, they involve the regulators in the product development process. It was reported that, when Sina was designing Weibo, it worked closely with the regulators (economist.com, 2011). In return, China's big Internet companies hold an almost monopoly position since foreign websites such as Facebook, Twitter, Instagram, Google have been blocked.

The relationship is likely to continue in the implementation of the Social Credit System. Government and Internet companies are likely to exploit BD in a way that is unimaginable and unacceptable in the West (FlorCruz, 2015). Instead of building its own system, the government may just wait for its private companies to build massive user databases and find a way to use them to monitor the population (Huang, 2015). This is a likely scenario since the Chinese political system provides the government with enormous capacity to intimidate the private sector.

It is also important to stress the importance of supportive informal institutions. In general, Asian societies seem to have a lower degree of concern about privacy. That is, violation of privacy is not yet viewed as a harm in Asian societies. A parallel can be drawn between the export of ZestFinance's algorithm and smart city technologies developed in the U.S. Asian's low degree of concern about privacy is arguably one of the main factors behind the development of South Korea's New Songdo City, which is described as the world's first greenfield smart city. An intriguing aspect of the development of the New Songdo City is that most of the core technologies were developed in the U.S. rather than in Korea (Kshetri et al., 2014). Supportive formal and informal institutions are the primary reason why U.S.-developed smart city technologies were first implemented in South Korea. Anthony Townsend, a research director Palo Alto, noted: "Much of this [smart city] technology was developed in U.S. research labs, but there are fewer social and regulatory obstacles to implementing them in Korea" (Regine, 2005). South Korea thus has an institutional advantage compared to other countries in the domains of smart city and the IoT. The New Songdo City makes an extensive use of RFID technology. For instance, using RFID, public recycling bins credit when a resident recycles a bottle (Regine, 2005). In general, while the use of RFID to automate tracking and monitoring people is becoming a serious concern in the West, privacy concerns are less visible in Asia (Rozek, 2007). Anthony Townsend, a research director Palo Alto, put the issue this way: "There is an historical expectation of less privacy. Korea is willing to put off the hard questions to take the early lead and set standards" (Regine, 2005). A similar point can be made about the export of U.S.-

developed algorithm to China. For instance, JD.com was reported to be the first customer of ZestFinance (Lohr, 2015).

As noted above, formal institutions do not guarantee to privacy. For instance, the website of Credit China allows anyone to check anyone else's credit score. Informal institutions are less concerned about privacy. For instance, for IoT deployment for the purpose of insurance underwriting or rating, whereas homeowners can voluntarily provide data related to household management in the West, Chinese consumers may lack awareness of their right to household management-related data and information.

It is also important to look at some of the fundamental driving forces of institutional changes that have facilitated the deployment of BD in the financial market. In the past, Banks' internal reward and punishment systems had an important impact on their lending behaviors. For instance, a failed loan to an SOE was less likely to be punished severely. On the other hand, a bank employee could lose job if he/she extends a loan to the private sector that is not paid off (Elliott et al., 2015). From a social and career standpoint, it was also important for a bank employee to make loans to entities run by powerful CCP members. The employees also used to face direct pressure and threats from CCP members or state officials. While to some extent, such pressures still exist today, their importance has decreased drastically (Elliott et al., 2015). This means that there is more incentive to engage in data-driven decisions.

Some changes may also hinder the use of BD in the financial market. For instance, there has been an increased public awareness and activism regarding the handling of personal information by BD companies. Misuse and abuse of personal data and information are reported to be key problems facing Chinese consumers who receive credits from the BD companies. Online credit services in China that use BD analytics in extending loans are being accused of abusing personal data to collect debts. One such example is Alibaba's Ant Check Later, which allows users to delay payments and pay in installments. An online user reported that he was contacted by Ant Check Later for information about his friend, who owed money to the payment service. Ant Financial Services reportedly said that the practice of contacting a borrower's friends or relatives to help with collecting debts is common in the financial sector. TMTpost cited a China Youth Daily poll, which showed that 75.9% of respondents believed there was abuse of BD (wantchinatimes.com, 2015). Users of JD.com reported similar problems. JD.com financial unit operates JD Baitiao, which is similar to Ant Check Later. In a question posted on online legal advice site 110.com, a JD.com user asked if it was legal for JD.com to give a third-party service his personal information for the purpose of debt collection.

7. Implications and concluding comments

The above discussion indicates that a main reason why low-income families and micro-enterprises in emerging economies such as China lack access to financial services is not because they lack creditworthiness but merely because banks and financial institutions lack data, information and capabilities to access the creditworthiness of this financial disadvantaged group. Thanks to BD, the financial industry has been undergoing a dramatic transformation that is likely to change the way financial services providers operate and the way they deliver their products and the way financial services are produced, distributed, and consumed. The use of BD can help

financial institutions overcome two major challenges discussed above: reducing information opacity and transaction costs. Overall BD has potential to improve SMEs' access to credit, in terms of both quantities and terms. BD-based models may also pose challenges to the current model of financial markets in China.

China's BD companies are faced with a number of contradictory and conflicting forces through which they must maneuver the economic, political and social terrains. Additional barriers may emerge when the share of the BD companies in the financial market increases. In addition to legitimacy from the government, securing legitimacy in the eyes of their customers is also important for China's BD companies. It is thus important to develop a more appropriate way of handling customer information.

Analysts disagree as to whether China's banks or ecommerce companies have a competitive advantage. One view is that companies such as Alibaba are likely to have a lead over banks such as the CCB because an ecommerce company is likely to be closer to small businesses and thus is likely to have better data. The opposite argument is that big banks have an advantage because they already possess a huge amount of data and information about clients (Rabinovitch, 2013). Irrespective of whether big banks or ecommerce companies have an advantage over the other group, the increasing prevalence of data driven decisions is likely to play a key role in expanding access to financial services in China.

Observers have noted that many practitioners lack skills and knowledge in various BD areas and they are not equipped to gain meaningful insights from diverse datasets. This situation is even more pronounced in the context of developing economies such as China. For example, if a bank wants to use data obtained from telecommunications operators, grocery stores, and utilities companies, the bank may need expertise and training in order to assess value and meaningfulness of each category of data, the appropriate level of details needed, and the most effective combinations of various categories of datasets (Baer et al., 2013). A related point is that some categories of data may be difficult to obtain. For instance, governments are likely to be cautious and hesitant about sharing citizens' identity and other information with private companies.

Reliability and validity issues have been adequately addressed in the processes used in assessing consumers' and businesses' creditworthiness. For instance, the use of factors such as reputation of friends, hobbies, time of the day the person is shopping online, types of items purchased in assessing individuals' creditworthiness may be of questionable validity and subject to a number of limitations. The government has a tendency to devote relatively fewer resources in relatively low priority areas such as increasing access to the credit market compared to other issues that it considers more urgent such as those related to maintaining national unity and social stability. This is especially important since state-owned banks account for the vast majority of the lending. This tendency is likely to be a huge barrier for the development of lending models that are based purely on economic fundamentals.

Finally, while BD has led to a number of innovations in consumer and business financing and the business models and approaches discussed above have widened the availability of financial services to more economically disadvantaged groups, they have not yet been able to do so for the

poorest of the poor. BD firms still lack reasonable amount of information on the poorest people, who do not yet have a cellphone or a social networking account.

References

ABC [Agricultural Bank of China] (2014). Annual report of ABC 2013. Beijing. <www.abchina.com/en/investor-relations/performance-reports/annualreports>.

Adams, D. W., & Nehman, I. G. (1979). Borrowing costs and the demand for rural credit. *The Journal of Development Studies*, 15(2), 165–176.

alibabagroup.com (2015). Ant financial unveils China's first credit-scoring system using online data. <http://www.alibabagroup.com/en/news/article?news=p150128>

Alois, J.D. MYBank, a Privately Owned Internet Bank in China, is Set to Challenge Traditional Banks, July 6, 2015, http://www.crowdfundinsider.com/2015/07/70665-mybank-a-privately-owned-internet-bank-in-china-is-set-tochallenge-traditional-banks/.

Asia Unhedged, June 25 (2015). Jack Ma's online bank lifts off, <http://atimes.com/2015/06/jack-mas-online-bank-lifts-off/>.

Baer, Tobias, Tony Goland, & Robert Schiff (2013). New credit-risk models for the unbanked, <http://www.mckinsey.com/insights/risk management/newcredit-risk models for the unbanked>.

Beck, Thorsten, Demirguc-Kunt, Asli, & Maksimovic, Vojislav. (2004). Bank competition and access to finance: international evidence. *Journal of Money, Credit, and Banking*, 36(3), 627–648.

Bernanke, & Ben, S. (2013). The crisis as a classic financial panic. In *The Fourteenth Jacques Polak Annual Research Conference* http://www.federalreserve.gov/newsevents/speech/bernanke20131108a.htm

Batra, G., Kaufmann, D., & Stibe, A. H. W. (2003). *Invest-ment Climate around the World: Voices of the Firms from the World Busi-ness Environment Survey*. Washington, D.C: World Bank.

BOCOM [Bank of Communications] (2014). Annual report of BOCOM 2013. Shanghai. <www.bankcomm.com/BankCommSite/en/invest relation/company develop.jsp?type=report>.

businesswire.com (2015). China rapid finance pre-approves 50 million Chinese consumers for first-time loans. <http://www.businesswire.com/news/home/20150218006682/en/China-Rapid-Finance-Pre-Approves-50-MillionChinese#.VTRqmDSlEg>.

Chaia, A., Goland, T., & Schiff, R. (2010). Counting the world's unbanked. McKinsey Quarterly, 2, 98–99.

Chen, Lulu Yilun September 23, 2014 Alibaba Arm to Create $163 Billion Loans Marketplace <http://www.bloomberg.com/news/articles/2014-09-23/alibabaarm-aims-to-create-163-billion-loans-marketplace>.

China Country Report. (2009). *Country forecast. Political risk yearbook: China country report, 2010*. East Syracuse, NY: The PRS Group.

chinacopyrightandmedia.wordpress.com (2015). Planning outline for the construction of a social credit system. <https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-asocial-credit-system-2014-2020/>.

Chng, J., Zhang, J. (2014). Big data: the next leading edge in the financial industry, <http://www.atkearney.com.tr/financial-institutions/ideas-insights/article/-/asset publisher/LCcgOeS4t85g/content/big-data-the-next-leading-edge-inthe-financial-industry/10192?101 INSTANCE LCcgOeS4t85gredirect=%2Ffinancial-institutions%2Fideas-insights%2F-%2Fassetpublisher%2FSxO47XYyZBF5%2Fcontent%2Fthe-new-reality-of-gccbanking%2F10192>.

Cui, E. 2013. Internet Firms Burst Into Banking, December 11, GKDragonomics.

Clarke, George, R. G., Robert Cull, Maria Soledad Martinez Peria (2001). Does foreign bank penetration reduce access to credit in developing countries? Evidence from asking borrowers, World Bank Policy Research Working Paper No. 2716.

Cull, R. & Martinez Peria, M.S. (2007). Foreign bank participation and crises in developing countries, World Bank Policy Research Working Paper Series, WPS4128.

Daokui Li, D., Feng, J., & Jiang, H. (2006). Institutional entrepreneurs. *American Economic Review*, 96(2), 358–362.

economist.com (2013). Giant reality-check four of the world's biggest lenders must face some nasty truths. <http://www.economist.com/news/finance-andeconomics/21584331-four-worlds-biggest-lenders-must-face-some-nastytruths-giant-reality-check>.

economist.com (2011). A nascent tech hub may succeed by solving local problems, April 4, 2015. <http://www.economist.com/news/business/21647624-nascent-tech-hub-may-succeed-solving-local-problems-techs-mex>.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: opportunities and challenges. *Academy of Management Journal*, 50(1), 25–32.

Elliott, D., Kroeber, A., Qiao, Y., 2015. Shadow banking in China: A primer, The Brookings Institution Hongyuran, W., Yuzhe, Z. and Ling, W. Online Financiers Banking on China's Future, 30 September 2015 http://english.caixin.com/2015-09-30/100859901.html.

FlorCruz, M. (2015). China to use big data to rate citizens in new 'social credit system'. <http://www.ibtimes.com/china-use-big-data-rate-citizens-newsocial-credit-system-1898711>.

forbes.com (2015). A beginner's guide to Alibaba group, <http://www.forbes.com/sites/joshsteimle/2015/01/26/a-beginners-guide-to-alibaba-group/>.

Golder, P. N. (2000). Historical method in marketing research with new evidence on long-term market share stability. *Journal of Marketing Research*, 37(2), 156–172.

Gottschalk, L. (1969). *Understanding history: a primer of historical method*. New York: Alfred A. Knopf.

Greene, D., & David, J. L. (1984). A research design for generalizing from multiple case studies. *Evaluation and Program Planning*, 7, 73–85.

GSMA (2015), July 13. China is global leader in deployment of internet of things, finds new GSMA report. <http://finance.yahoo.com/news/china-global-leaderdeployment-internet-062400669.html>.

Honohan, P. & King, M. (2009). Cause and effect of financial access: cross-country evidence from the Finscope surveys. World Bank Mimeo.

Huang, Z. (2015). October 9. All Chinese citizens now have a score based on how well we live, and mine sucks. <http://qz.com/519737/all-chinese-citizensnow-have-a-score-based-on-how-well-we-live-and-mine-sucks/>.

IDC [International Data Corporation] (2012). Chinese big data technology and services market forecast and analysis 2012–2016 [in Mandarin]. <www.idc.com.cn/uploadpic/5-Frank-banking2013.pdf>.

Jack, W. & Suri, T. (2010). The economics of M-PESA. <www.mit.edu/~tavneetMPESA.pdf>.

Jappelli, T., & Marco, P. (2002). Information sharing in credit markets: international evidence. *Journal of Banking and Finance*, 26, 2023–2054.

Jia, C. (2009). No vacancies on horizon for 12 m job seekers. China Daily. Retrieved from <http://www.chinadaily.com.cn/bizchina/2009-08/22/content 8602908.htm>.

Jing, M. (2014). Alibaba, lenders team up for SME financing, China Daily, July 23, <http://www.chinadaily.com.cn/business/2014-07/23/content 17901230.htm>.

Klein, B. P., & Cukier, K. N. (2009). Tamed tigers, distressed dragon. *Foreign Affairs*, 88(4), 8–16.

Koo, R. C. (2009). An American lost decade? Yes and no. *International Economy*, 23(3), 6–22.

Kshetri, N. (2014). *Global entrepreneurship: environment and strategy*. New York: Routledge.

Kshetri, N., Alcantara, L. L., & Park, Y. (2014). Development of a smart city and its adoption and acceptance: the case of new Songdo. *Communications & Strategies*, 96, 113.

Kshetri, Nir (2011). The global financial crisis: evidence from China and India, Thunderbird International Business Review, Special issue on the 2008 Financial Crisis, 53(2), March/April, pp. 247–262.

Li, J. P., Zhang, Y. J., Wu, D. S., & Zhang, W. (2014). Impacts of big data in the Chinese financial industry. The Bridge, 44(4) https://www.nae.edu/Publications/Bridge/128772/129193.aspx

Li, X.X. (2014). Alibaba has big hopes for new big data processing service. <http://english.caixin.com/2014-07-17/100705224.html>.

Lin, H. (2011). Foreign bank entry and firms' access to bank credit: evidence from China. *Journal of Banking and Finance*, 35(4), 1000–1010.

Mason, R. O., McKenney, J. L., & Copeland, D. G. (1997). An historical method for MIS research: steps and assumptions. *MIS Quarterly*, 21(3), 307–320.

McKay, C. & Pickens, M. (2010). Branchless banking pricing analysis, consultative group to assist the poor (CGAP). Available from: <www.cgap.org/gm/document-1.9.43876Branchless Banking Pricing Analysis May 2010.pdf>.

Micco, A., Panizza, U., & Yanez, ˜ M. (2007). Bank ownership and performance: does politics matter? *Journal of Banking and Finance*, 31, 219–241.

Mu, X. (2009). Soros: China will emerge as winner from current economic turmoil. <http://news.xinhuanet.com/english/2009-10/31/content 12363369.htm>.

Platt, G. (2010). Bankers get ready for rebound in world trade. *Global Finance*, 24(2), 29–36.

Regine (2005). Korea's U-city. Retrieved from <http://we-make-money-not-art.com/archives/2005/10/public-recyclin.php>.

Rogaly, B. (1996). Micro-finance evangelism, 'destitute women', and the hard selling of a new anti-poverty formula. *Development in Practice*, 6(2), 100–102.

Shu, C. (2015). Lending platform China rapid finance raises $35 M series C at a $1 B valuation, July 29, <http://techcrunch.com/2015/07/29/china-rapid-finance/>.

Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *American Economic Review*, 71(3), 393–410.

UNCTAD (2000). Building confidence: electronic commerce and development, United Nations Conference on Trade and Development, UNCTAD/SDTE/MISC.11.

Vidaillet, T. (2010, January 28). China says to maintain stable economic policies. Retrieved from: <http://www.reuters.com/article/idUSTRE60R4ZV20100128>.

volkskrant.nl. (2015). China rates its own citizens—including online behavior. Available from: <http://www.volkskrant.nl/buitenland/china-rates-its-owncitizens-including-online-behaviour~a3979668/>.

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14(4), 490–495.

Wildau, G. (2015). January 28. Alibaba's finance arm launches credit scoring service, <http://www.ft.com/intl/cms/s/0/34e77fe8-a6a3-11e4-9bd3-00144feab7de.html>.

Williamson, O. E. (1989). Transaction cost economics. *Handbook of Industrial Organization*, 1(3), 135–182, 1st ed.

Xinhua (2010). March 2. Economists say China may 'adjust' monetary stimulus. Retrieved from: <http://www.chinadaily.com.cn/bizchina/2010-03/02/content 9524080.htm>.

Xinhua (2015). June 2. China establishes credit score website. Retrieved from: <http://www.china.org.cn/china/2015-06-02/content 35716132.htm>.

Zakaria, F. (2009). The recession's real winner. *Newsweek*, 154(17), 18.

Zhou, X.Y. (2015). A small bank makes a big difference 2015. Retrieved from: <http://www.bjreview.com.cn/quotes/txt/2015-04-01/content 663611.htm>.