United Nations
Office for South-South Cooperation

South-South Ideas

Big Data's Role in Broadening Financial Inclusion in the Global South

United Nations Development Programme
304 East 45th Street
New York, NY 10017 USA

Disclaimer
The views expressed in this publication are those of the authors and do not necessarily represent those of the United Nations, UNDP or the United Nations Member States. The designations employed and the presentation of materials on maps do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations or UNDP concerning the legal status of any country, territory, city or area or its authorities, or concerning the delimitation of its frontiers or boundaries.

South-South Ideas

Big Data's Role in Broadening Financial Inclusion
in the Global South

February 2021

# Table of Contents

Acknowledgments

Executive Summary

Consumer and business finance companies and microcredit organizations have had limited success in serving the needs of economically-active low-income families and micro-enterprises in Southern countries. 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. Important technological and strategic driving forces are behind the implementation of Big Data techniques for such purposes. The objective of this paper is to examine this issue from the perspective of Southern economies. The approach of this study can be described as theory building based on multiple case studies. The study analyses seven cases of financial technology companies operating in the Global South.

The paper 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. The paper looks at various categories of personal financial and non-financial information being used as proxy measures for a potential borrower's identity, ability to repay and willingness to repay. The analysis of the paper indicates that the main reason why low-income families and micro-enterprises in 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 this financially disadvantaged group.

Introduction

PricewaterhouseCoopers' chief economist in Lagos, Nigeria, Andrew Nevin commented in regards to the future of lending: "In ten years' time, the technology will be so good that when someone applies for a loan, the only question you will ask is: 'Can we access your data?'" (Norbrook and Anderson, 2016). This vision is increasingly becoming a reality.

A Myanmar-based microfinance institution, Maha Agriculture, started combining harvesting data based on weather monitoring with its credit-scoring model which is expected to improve predictive capabilities and increase its borrowers (Bary, 2018).

According to the H2 Ventures/KPMG annual report *FinTech 100*, as of 2019, the top 100 financial technology (FinTech) firms, ranked on innovation, capital raising activity, size, and location, had raised over US$70 billion in venture capital and over $18 billion in other forms of capital (H2 Ventures and KPMG, 2019). A growing number of traditional banks are collaborating with FinTechs to provide innovative products and services to customers, whether it is through partnerships or alternative forms, such as digital platforms, open banking, Banking-as-a-Service, alliances and consortia (Shevlin, 2019).
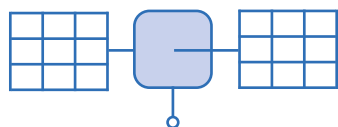
FinTechs began to fully utilize Big Data and machine learning capabilities to explore factors that have not previously been thought about or regarded as important determinants of the likelihood that customers will repay a loan (Delgado, 2016). For instance, a Myanmar-based microfinance institution, Maha Agriculture, started combining harvesting data based on weather monitoring with its credit-scoring model which is expected to improve predictive capabilities and increase its borrowers (Bary, 2018). In the near future, it will be possible to use satellite imagery of the roof of a farmer's house and the size of a plot for agricultural land for identification purposes and to assess the ability of a client to repay a loan. With the proliferation of low-cost satellites, it is becoming increasingly easy and affordable to get data and information to gain a better understanding of various economic, social and environmental indicators (Kshetri, 2016b).

FinTechs are also developing and implementing emerging technologies, such as artificial intelligence and machine learning, to assess creditworthiness and reduce fraud. For instance, China's Alibaba and Mexico's Kueski use artificial intelligence to tackle fraudulent activities in loan application process (López, 2020; Perez and Soo, 2017). Malaysia's Poladrone was reported to be developing artificial intelligence algorithms that process farms' aerial images to assess crop performance which would provide data to predict farmers' ability to repay (Voutier, 2019). similarly, Singapore's Adatos processes satellite imagery of farmland to produce harvesting data. Among Adatos' clients is a Thai company that accounts for 45 percent of farm machinery in Southeast Asia. Adatos was reported to be helping the company develop new precision agriculture products (The, 2018). All these are encouraging developments because consumer and business finance companies and microcredit organizations have had limited success in serving cost-effectively and sustainably the needs of economically-active low-income families and micro-enterprises.

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: inefficiency and informational opacity. First, traditional banks have been unwilling and reluctant to serve small-scale borrowers, such as low-income people and small businesses, due to high transaction costs and inefficient processes associated with making small loans to these borrowers (Adams and Nehman 1979; Rogaly, 1996). Secondly, informational opacity has been a barrier to access (Stiglitz and Weiss, 1981). Traditionally, a national credit bureau collects and distributes reliable credit information, hence increasing transparency and minimizing a bank's lending risks. Many Southern economies are characterized by the lack of, or poor performance of, credit rating agencies to provide information about the creditworthiness of small and medium enterprises (SMEs). This puts SMEs at a disadvantage in the credit market as SMEs tend to be more informationally opaque than large corporations because the former often lack certified audited financial statements. Thus, it is difficult for banks to assess or monitor the financial conditions of SMEs (Kshetri, 2019).

Different lending behaviours of different groups of banks in terms of the propensity to lend to low-income people and small businesses can also be explained in terms of access to information. For instance, Beck et al. (2004) found that domestic banks had a 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. This finding underscores the need for data and information in making sound lending decisions. 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 problems related to lack of access to financial services can be largely eliminated by creating better risk models using increased computing power and new sources of data and information (Baer et al., 2013).

A key mechanism underlying this transformation concerns the use of Big Data in assessing, evaluating and refining the creditworthiness of potential borrowers and reducing transaction costs. Scholars and practitioners have suggested that healthcare, finance and insurance are industries that are likely to be transformed by artificial intelligence, machine learning and Big Data (Kshetri, 2016b; Marr, 2016). Possible data sources include social media and mobile phone usage patterns and utility bill payment history. The Fair Isaac Corp (FICO), the creator of the most widely used credit scores in the United States, has started using nonfinancial/non-payment data to determine scores, replacing traditional FICO scores when loan or bill payment information is lacking.

Many successful examples of utilizing Big Data to expand financial services in Southern economies exist. Various FinTech companies have taken advantage of Big Data to build alternative credit scores for low-income individuals (World Bank, 2019). These companies, such as Aynnah in the Philippines, CredoLab in Indonesia, Malaysia, Singaporeand many other countries, allow customers to use apps to apply for loans and get an immediate determination of whether to extend the loan. The app scans the digital footprint of a mobile phone which includes their contacts, social network profiles and geographical patterns. Such automated credit scoring provides flexibility when evaluating financial trustworthiness and contributes to expanding credit.

Big Data has a major role in financing SMEs. SMEs often face difficulties getting loan approvals from traditional financial institutions because their risk is harder to assess due to the lack of long periods in business and a dearth of public information. An example of a response to that problem, Alibaba, one of the largest online retailers in the world, compiles data on their monthly sales and preapproves loans and provides credit to SMEs through its own digital bank called MYbank. Alibaba's unique '310 model' enables borrowyuaners to complete online loan applications in three minutes and obtain approval in one second — with zero human intervention. MYbank has provided services to roughly half of the millions of SMEs in Chin (Chataing and Kushnir, 2018).

Big Data promotes financial networks for underserved groups. In several emerging economies, mobile money still depends on networks of agents who helps customers deposit and withdraw cash. As such, it is important to ensure that agents have sufficient money and are located close enough to their customers. Zoona, a financial services provider in Zambia, uses simulations, data from field staff and Google Maps to determine the optimal location of agents and the potential demand (Mastercard Foundation and IFC, 2017). DBS Bank, a Singaporean bank, uses data on ATM withdrawals and deposits to forecast ATM cash demands and optimize its ATM network (Fitzgerald, 2014).

A final issue that deserves mentioning relates to the fact that like any other economic phenomenon (Parto, 2005), Big Data company operations have important institutional components and implications. Institutions are "macro-level rules of the game" (North, 1990, p. 27), which include: a) formal institutions, such as rules, laws, constitutions; and b) informal institutions, such as social norms, conventions and self-imposed codes (North, 1996). It is not clear how formal and informal institutions enable the operations of Big Data companies in the financial sector of Southern economies. This points to potential negative implications when formal and informal institutions operate Big Data poorly.

The objective of this paper is to look at the role of Big Data and associated context in facilitating the access to financial products for economically-active low-income families and micro-enterprises in emerging economies. Specifically, the paper aims to address the following research questions:

Research question 1 (RQ1)
How can Big Data help reduce transaction costs and information opacity and thus address barriers to financial services for disadvantaged groups in the Global South, particularly low-income families and microenterprise owners and operators?

Research question 2 (RQ2)
How can different dimensions of Big Data address barriers to financial services for disadvantaged groups in the Global South?

Research question 3 (RQ3)
What roles should there be for formal and informal institutions to facilitate the operations of Big Data companies in the financial and banking sectors of the Global South?

Research question 4 (RQ4)
How is South-South cooperation being leveraged to expand the impact of Big Data on financial inclusion?

The paper is structured as follows: a review of the relevant literature; a discussion of the method employed in this study and a brief overview of the selected cases; observations from the selected cases to examine Big Data as a tool for increasing access to financial services for disadvantaged groups. Finally, there is a section on discussion and implications and concluding comments.

1. Literature Review

The literature review was structured around three key aspects of this study: a) barriers to 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.

1.1. Barriers to Access to Financial Services

Women in Southern economies are nine percentage points less likely than men to have a bank account. In particular, the Middle East and North Africa regions show the largest gap, with 52 percent of men and only 35 percent of women having an account.

Access to financial services has been improving globally, but large disparities remain between the rich and poor and between men and women, as well as across countries. According to the 2017 Global Findex database, the share of adults who have an account with a financial institution or through a mobile money service in Southern economies rose from 54 percent to 63 percent between 2014 and 2017 (World Bank, 2018). This is still far below the global share, which is from 62 percent to 69 percent. Furthermore, women in Southern economies are nine percentage points less likely than men to have a bank account. In particular, the Middle East and North Africa regions show the largest gap, with 52 percent of men and only 35 percent of women having an account. In addition, financial institutions still find SME financing challenging due to credit risks, the macroeconomic environment, government policy changes, technology adoption, client engagement and competition (IFC, 2019).

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 Southern economies since the mid-1990s (Cull and Martinez, 2007). A study of banking sector assets in 104 Southern countries indicated that during 1995-2002, the average share held by foreign banks increased from 18 percent to 33 percent (Micco et al., 2007). This trend is accelerated by globalization, financial integration, and greater regionalism with the increase of South-South cooperation.

However, large foreign banks tend to be reluctant to finance SMEs (Clark et al., 2001). Furthermore, foreign banks cherry-pick borrowers (Azmeh, 2018) which prevents overall access to the financial sector (Azmeh, Al Samman, and Mouselli, 2017; Beck and Martinez, 2008; Detragiache, Tressel and Gupta, 2008; Gormley, 2010). Large foreign banks tend to prefer lending to hard information borrowers over soft information borrowers because of cultural and geographical distance between their headquarters and local branches and easier access to external liquidity from their parent banks (Mian, 2003, 2006). Such cherry picking worsens the credit pool and eventually excludes soft information borrowers from borrowing. The negative effect of cherry picking is more pronounced in Southern economies where the relationship-based lending is important.

Some initiatives have been taken to deal with the limitations of general consumer and business finance companies. For instance, microfinance started in the 1970s to provide small working capital loans to low-income people in the Global South to start a business. A microfinance institution typically borrows funds at a low cost and tries

to keep loan defaults and overhead expenses very low. Loans are made to entrepreneurs without physical collateral. However, demand exceeds supply by a large factor in the global microfinance industry and market (Kshetri, 2014). A 2007 report from the Deutsche Bank indicated that due to the funding limitations of microfinance institutions, only 10 percent of potential borrowers get loans. It was reported that one microfinance institution in India had a waiting list of 50,000 clients that were seeking but unable to find loans (Engen, 2009).

The loan portfolios of most microfinance institutions in Southern countries are typically concentrated in urban areas. Systemic risks associated with droughts, floods, cyclones and other extreme weather-related events tend to make agricultural loans less attractive and hinder the ability and enthusiasm of these institutions to expand their services to rural farmers (Miranda and Gonzalez-Vega, 2011).

Crowdfunding, a newer form of entrepreneurial financing, has exhibited a high degree of West centricity. For instance, in 2012, per capita crowdfunding investments were about $3 in North America, $1.30 in Europe and $0.02 in the rest of the world (Kshetri, 2015). Meanwhile funds raised by region in 2018 are as follows: a) South America - $85.7 million; b) Africa - $24.1 million; and c) Asia - $10.5 billion (Startups, 2018). Crowdfunding in Asia has increased over the past few years, while the West has been unable to fulfill orders promised to backers. The lack of regulations and oversight agencies that vet crowdfunding ventures in emerging economies have delayed potential innovation in crowdfunding.

1.2. The Transaction Cost Economics Approach

The transaction cost economics approach (Williamson, 1989) can provide insights into severe barriers faced by small businesses and low-income populations in accessing financial products and services from conventional financial institutions. Traditional banks are unwilling to serve small-scale borrowers because 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, borrowers' 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 to reduce the transaction costs of making small loans to low-income people and small businesses (Rogaly, 1996).

$0.01

Average cost of a banking transaction on the Internet.

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 $1.27 in a branch and $0.27 in an ATM, whereas it is $0.01 on the Internet (UNCTAD, 2000). A study indicated that, for a transaction involving $23, branchless banks cost 38 percent less than commercial banks and 54 percent less than informal money transfer channels (McKay and Pickens, 2010). For instance, the average mobile transaction conducted via Kenya's mobile phone-based money transfer, financing and microfinancing service M-PESA is about a hundredth of the

average check transaction and half of the average ATM transaction (Jack and Suri, 2010). Time saving is a major benefit of the M-PESA system. It is estimated that each M-PESA transaction requires two-three hours of time and costs of $3 if it is processed through the traditional financial channel (Schwartz, 2014). Likewise, M-Shwari, a bank account offering a combination of savings and loans offered jointly by Safaricom and the Commercial Bank of Africa in Kenya, allows customers to use cellphones to deposit savings into "locked" accounts. The savings can be unlocked on a specific date, such as the day when school fees for children need to be paid (Economist.com, 2014).

1.3. Informational Opacity, Moral Hazard And Adverse Selection

The barriers faced by disadvantaged groups to access financial products from conventional financial institutions often result from informational opacity, which may lead to moral hazard and adverse selection problems (Stiglitz and Weiss, 1981).[1] For instance, it is possible that a borrower has provided false and misleading information to a lender about their assets, skills and credit capacity, but the lender provided a loan due to the inability to verify the information. This is a problem of adverse selection associated with the information opacity related to ability to repay loans. Likewise, the borrower may have a capacity to pay but may not have entered into a contract with the lender in good faith. This is a problem of moral hazard due to the information opacity related to willingness to repay loans.

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 Pagano (2002) provided empirical evidence which suggested that the degree of information sharing between intermediaries is positively related to financial development. Availability of information can help fight moral hazard and adverse selection problems.

---

1   Adverse selection (anti-selection or negative selection) arises from information asymmetry between buyers and sellers prior to a deal. In such a case, one party is unable to determine if the other party is lying. Likewise, moral hazard is the problem of not being able to determine if the other party is cheating or acting dishonestly following a deal.

2. Method
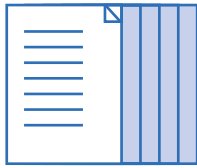
The approach of this study can be described as theory building from multiple case studies.

The study explores approaches of theory building from multiple case studies, which is becoming increasingly popular in social science (Eisenhardt and Graebner, 2007; Kshetri, 2016b). Compared to a single case study, multiple case studies provide a stronger base for theory building (Rowley, 2002; Yin, 1994). Connection with related literature, establishment of theoretical gaps that exist in the literature and an explicit statement of research questions to address the gaps are the key features of strong empirical research (Eisenhardt and Graebner, 2007). In qualitative research, it is also important to make a strong case for the importance of the research questions that have been raised (Bansal and Corley, 2012). For this study, the theoretical and practical importance of research on the roles of Big Data in enhancing accessibility and affordability of finance which can lead to poverty alleviation, are established.

There has been a good deal of debate on whether case research should be based on theory specified a priority or on grounded theory. It is thought that Big Data enhances accessibility and affordability of finance and may lead to poverty alleviation. Whyte (1984) argues that, to be valuable, research should be guided by "good ideas about how to focus the study and analyze those data'' (p. 225). On the contrary, Glaser and Strauss (1967) suggested that evolution of a theory from the data is the basis for development of grounded theory rather than an imposition of *a priori* theory. Likewise, Van Maanen, Dabbs and Faulkner (1982, p. 16) suggested that investigators avoid prior commitment to any theory. In this study, we follow Whyte's approach.

2.1.   Selection of Cases

Broadly speaking the selection of cases in multiple case study research has the same objectives as in random sampling. That is, the cases should represent the population and a useful variation on the dimensions of theoretical interest is needed (Seawright and Gerring, 2008). A key difference is that in a multiple case study design, the choice of cases needs to be made more on a substantive than statistical basis to adequately represent a target population (Greene and David, 1984).

First, it is important to make it clear that case selection is also guided by pragmatic, logistical and financial reasons (Seawright and Gerring, 2008). Cases were selected for which it was possible to obtain sufficient information from secondary resources. Eisenhardt (1989) suggested that about seven cases would be ideal for building theory. Following this recommendation, we selected seven cases of FinTech companies operating in the Global South. To select the cases, two methods were combined: extreme case method and diverse case method (Seawright and Gerring, 2008). More specifically, the process started with extreme case method and morphed over time with implementation of different requirements and recommendations.

In the extreme case method, cases with extreme values on the independent (X) or dependent (Y) variable of interest are selected (Seawright and Gerring, 2008). The FinTech companies selected in this paper are extreme in the sense that they are among the most successful in the Global South. That is, we did not choose any unsuccessful or average FinTech companies.

Seawright and Gerring (2008) suggest that if the researcher has some idea about additional factors that might have effect on Y (the outcome of interest), it would be better to pursue other case selection methods. Following this recommendation, a diverse case method was utilized as a strategy to select specific cases of successful FinTech companies with diverse characteristics. A key objective in this method was to achieve maximum variance along relevant dimensions (Seawright and Gerring, 2008). This method requires the selection of two or more cases to represent the full range of values characterizing X, Y or some relationship between these variables (Seawright and Gerring, 2008).

As to the factors affecting Y, especially researchers (Kshetri, 2016b) and practitioners (Norbrook and Anderson, 2016) have emphasized the importance of high-quality data. Moreover, Big Data is as much about analytics as it is about the data. The improvements in analytical capacities (e.g., statistical machine learning and algorithms) are the key *to* understanding patterns and trends in data (scidev.net, 2014). To achieve diversity, cases with different levels of data and analytical capabilities were selected.

The cases selected in this study include Big Data operating in diverse geographic areas (e.g., China, Colombia, Kenya, Mexico, the Philippines and Uganda), focusing on diverse groups of target customers (small holder farmers, small businesses and individuals who lacked access to financial services in the non-Big Data environment). The countries in which FinTech companies selected for this study operate are characterized by different formal and informal institutional structures. For instance, the degree of institutionalization of data security and privacy varies across the above countries. Omidyar Network conducted an in-depth study of the early adopters of what they refer to as "Big Data, Small Credit" in Colombia and Kenya . Big Data, Small Credit involves utilization of Big Data to assess the creditworthiness of potential borrowers. Among Colombian consumers who borrowed from Lenddo, about 70 percent were willing to share information about social media activity and web browsing history with the lender to improve the chance of getting a loan or getting a larger loan. In Kenya, only five percent of respondents were reported to feel uncomfortable to provide share information with a financial institution (Costa et al., 2015).

A further way to increase diversity would be to include cases that have various causal paths that link the related variables to a particular outcome. For instance, three different independent variables (X_1, X_2 and X_3) may have an effect on Y, but they may do so independently and in different ways (Seawright and Gerring, 2008). In addition to the quality of data and analytical capabilities (X_1), two additional independent variables — financial institution reliance on Big Data organizations for analytical capabilities and the location of data (external vs. internal) — emerged as having effects on the dependent variable. The case selection process thus morphed as the research moved forward and a better understanding emerged of the functioning of FinTech companies.

The cases selected in this study include Big Data operating in diverse geographic areas (e.g., China, Colombia, Kenya, Mexico, the Philippines and Uganda), focusing on diverse groups of target customers (small holder farmers, small businesses and individuals who lacked access to financial services in the non-Big Data environment).

Table 1 presents the cases selected and business models used by the selected Big Data companies. The table classifies the selected cases in terms of financial institution reliance on Big Data organizations for analytical capabilities and the dominant sources of data used (external vs. internal). Three replicated cases have been used in which a Big Data organization is also a provider of financial services. While both Kilimo Salama and Alibaba mostly use internal data, Kueski relies on external data. These replicated cases have high and low degrees of different attributes of interest. Four replicated cases of financial institutions have been used which rely on Big Data organization expertise and experience. While Agrilife mostly uses internal data, Cignifi, Lenddo and Kreditech rely on external data.

Table 1: *Business models used by Big Data companies to enhance access to finance.*

Loans/financial services provided by	Dominant sources of data used by the processor of big data	
	Internal	**External**
Big Data organization (processor of Big Data and provider of financial services as the same organization)	[I] **Kilimo Salama** **Alibaba**	[II] **Kueski**
Financial institutions relying on Big Data organizations' expertise and experience	[III] **Agrilife**	[IV] **Cignifi** **Lenddo** **Kreditech**

Sources and characteristics of data

Prior researchers have identified various dimensions of data quality, which are central to obtaining valid and reliable results (Golder, 2000; Gottschalk, 1969; Jayawardene et al., 2013; Loshin, 2001; Mason et al., 1997). For instance, Gottschalk (1969) suggested that the sources of evidence, as well as the evidence, need to be evaluated using criteria, 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. Regarding the last point, previous researchers have recommended that data and information be triangulated from multiple sources (Constantino and Westberg, 2009).

First, it is important to make it clear that this study mainly relies on archival data, which is among a variety of recognized data sources for case studies (Eisenhardt and Graebner, 2007). For instance, Ansari et al. (2015) mainly relied on archival data sources to examine how the digital video recorder TiVo dealt with the disruptor's dilemma by adjusting its strategy, technology platform and relation with various players in the TV industry ecosystem. As noted above, this choice is due to pragmatic, logistical and financial reasons (Seawright and Gerring, 2008) as well as considerations related to ease with which data can be located and gathered (Stvilia et al., 2007; Wang and Strong, 1996).

For this study, attempts were made to assess the coherence and internal consistency of the data. Following the suggestion of prior researchers (HIQA 2011), coherence was evaluated by comparing different data items for the same point in time and the same data items for different points in time. For instance, for the numbers of farmers insured by Kilimo Salama (Case 2), data from two sources were matched for 2013 (ifc.org, 2016; Kalan, 2013) and for 2014 (Acre Africa, 2015; businessgreen.com, 2016). Likewise, data were compared for 2011, 2013 and 2014 to ensure there were logical patterns.

A key dimension of data quality is reputation and trustworthiness. This means making sure that the source as well as the content of the data are trustworthy (Wang and Strong, 1996). A related characteristic is objectivity. That is, the data are unbiased and impartial (Wang and Strong, 1996). This is related to accuracy or correctness. The goal is to make sure that the information is free from distortion and bias (Eppler, 2006). Another key point that must be considered is an accurate mapping of the real-world phenomenon (Price and Shanks, 2005). To achieve these various goals, among other things, as noted above, for this report the data and information were corroborated from multiple sources. Also, the authors mostly relied on information reported by reputable third parties instead of pulling descriptions directly from the websites of the organizations chosen. These steps minimized potential self-reporting biases. Another consideration is the timeliness and currency of the data (Wang and Strong, 1996). Case study researchers must make sure that information is up-to-date and not obsolete (Eppler, 2006). In this regard, the authors insured that the age of the data was appropriate to study the cases selected. The authors also followed the latest news related to the cases chosen. In addition, they viewed the websites of the relevant companies.

2.2.   Case studies

In this section, the FinTech firms selected in this study are briefly described.

Case 1

Agrilife

Agrilife is a cloud-mobile platform of the Kenya-based mobile payment solution and service provider MobiPay. Agrilife connects farmers with value chain partners, such as dairy processors who purchase milk, credit appraisers and local input/agro-dealers. Agrilife is seen as a "one-stop-virtual-agri-info-shop" (cta.int, 2014) and provides financial institutions and suppliers "near-real-time information" on farmers' ability to pay for services (capacity.org, 2013). Agrilife reduces transaction costs by linking the various value chain actors. A farmer can make credit requests via a mobile phone. The credit appraiser uses a range of data to assess the creditworthiness. The input provider makes decisions on credit.

In the past, a farmer's needs may have been known by a field officer or a programme manager of a development organization only if the farmer was part of a household survey. Survey data take a relatively long time to analyze. By analyzing data of many farmers through mobile money and other types of transactions, a credit appraiser assesses farmer creditworthiness. Based on the milk sold to a dairy processor, for instance, a farmer receives an SMS short code which can be used to access credit to purchase inputs. In this way, the farmer can use the future production of milk as collateral. When a farmer sells milk, a Bluetooth enabled digital scale is used.

The transaction's weight details are transferred to the platform, which minimizes data discrepancies (cta.int, 2014). Digitization therefore minimizes the governing or transaction costs associated with opportunistic behaviours and uncertainty (Rindfleisch and Heide, 1997; Williamson, 1979, 1985).

As of September 2013, Agrilife facilitated over $2 million in revolving credit lines to about 120,000 small farmers in Kenya and Uganda. Century Microfinance Bank, while using the Agrilife platform, has increased its outstanding loan portfolio from KSH 25.2 million to KSH 88.6 million. Century could increase its loan portfolio by 250 percent in five months with minimal extra costs (Grossman and Michael, 2014).

Kilimo Salama

Case 2 →

The social enterprise Kilimo Salama (which means "safe agriculture" in Swahili) is a highly visible example of an organization involved in the development and use of index insurance to serve a vulnerable market that traditional insurance schemes do not reach. Kilimo Salama is a partnership between the Syngenta Foundation for Sustainable Agriculture, UAP Insurance and Safaricom. It has developed a micro insurance scheme for small farmers in Kenya to protect them against poor weather conditions. Its weather-based index insurance uses data collected by weather stations across the country to pay to the policyholders in case of bad weather. Farmers can buy the insurance at the beginning of the season for about 10-20 percent of the amount they invest in seeds and inputs (Kalan, 2013). For some farmers, the cost of insurance amounts to as little as 1 kg of maize, seed or fertilizer (un.org, 2014). The insurance, which is completely automated, is distributed through dealers who use an advanced phone application with camera and phone functions to scan and capture policy information through a code. The information is uploaded to Safaricom's mobile cloud-based server, which administers policies. Farmers receive information about their policy and payouts in SMS messages (Schneider, 2013). In 2011, when there was insufficient rainfall in their areas, 1,400 farmers received payouts (african-businessreview.co.za, 2020).

Weather stations are equipped with wireless sim cards that transmit data every five minutes to a cloud-based server. At the end of every season, the data is aggregated and combined with satellite data to map out rain patterns. Kilimo Salama works with agronomists to calculate the index and identifies the locations that experienced too much rain, too little rain or rain at the wrong time. Farmer payouts are calculated based on crops, location and the amount invested in seeds (Kalan, 2013). Another benefit to farmers is that banks and microfinance institutions are more comfortable giving loans to farmers thanks to the insurance scheme. In this way, access to essential credit is becoming easier for farmers. Kilimo Salama started in 2009 as a small initiative with only 200 farmers; by 2014, the now-enterprise covered close to 200,000 farmers in Kenya, Rwanda and Tanzania with a total sum insured of $12.3 million (Sibiko et. al., 2018; Greatrex et.al., 2015).

Case 3

Cignifi

Massachusetts -based Cignifi provides credit scores for people who lack traditional credit histories, which helps them get access to financial services. The company mainly relies on mobile phone behavioural data. A technology developed by the company can recognize patterns in consumer phone-calls, text messages and data usage, which are used to predict lifestyle and credit risk profile (bigdata-startups.com, 2013). As of 2014, the company was working with mobile carriers and financial institutions in emerging markets, such as Brazil, Chile, Ghana and Mexico. The company had offices in Accra, Mexico City and São Paulo. (Kokalitcheva, 2014). In 2016, Cignifi partnered with Equifax, one of the three largest credit-reporting firms in the U.S. in terms of the number of individuals and businesses covered, to help Equifax expand its credit scoring capabilities in Latin America. The two firms are working to determine consumer credit risk based on cell phone usage data in Latin America (Andriotis and Demos, 2016).

Case 4

Kueski

In Mexico, banks are extremely risk averse. In 2017, approximately 9.52 percent of Mexicans owned a credit card, down from over 17.82 percent in 2014 (Statista, 2020). While they made up more than 99.8 percent of the country's enterprises in 2014, micro-, small and medium-sized firms in Mexico received only 11.1 percent of total bank credit for businesses, including loans, overdrafts and credit cards (*Financial Times*, 2017). Most Mexicans also lack a meaningful credit score, which has hindered the ability of traditional Mexican banks to provide them with consumer loans. Guadalajara-based Kueski provides loan services to Mexicans without any collateral or a personal meeting with the borrower. Kueski uses alternative information, such as the number of social media friends, to provide faster and easier credit. Individuals can apply for a loan with Kueski via a computer or a smart phone.

Unlike some other Big Data companies, such as Cignifi, that work with mobile carriers and financial institutions, Kueski itself is a loan service provider. The company's Big Data team employ machine learning methods for risk assessment using an algorithm and adaptation over time (Flores, 2016). It analyses credit history, information collected using a credit questionnaire, online activity, socio-demographic information and other relevant data. The cloud computing firm Mambu provides Kueski with technology and computing power, such as a customer relationship management platform, accounting and systems integration. Mambu also helps Kueski connect its payment gateways with banks (Holley, 2014.) Promising customers receive loans averaging $150 for periods which average 22 days (Economist.com, 2015). As of July 2019, Kueski provided more than 1.5 million loans worth $220 million (Zerucha, 2019). It also makes use of artificial intelligence, especially for fraud prevention. Data from various sources are used for this purpose (López, 2020).

Monedo

Monedo (rebranded from Kreditech in early 2020) is a German technology company that uses an online platform to assess loan applicant creditworthiness. Monedo was reported to process 15,000 real-time data points, including location, social media activity and e-commerce behaviour, to generate a credit score for a consumer. It then sells the score to retailers. The company's risk model does not require an external credit bureau to conduct identification, fraud detection and scoring decisions (Wicem, 2014). As of 2015, Monedo (then Kreditech) had served more than 800,000 customers who lacked the credit history to borrow money through other means and issued over two million loans in countries like India, Poland, Spain and Russia, (Zendesk, 2017). As of March 2020, Monedo had more than one million customers who had received over two million loans by then (Alois, 2020). In 2018, Monedo issued about $120 million of credit (Kreditech, 2019).

Lenddo

Lenddo, a Singapore-based software-as-a-service company, uses social media and smartphone records to build credit worthiness and get access to financial services. As of mid-2014, Lenddo's network had about one million members who gave Lenddo access to all of their social media activities from Twitter, Facebook, LinkedIn, Google, Yahoo and Hotmail. It was reported that a typical Lenndo credit application had over 12,000 data points that could be used to assess creditworthiness (aws.amazon. com, 2015). Lenddo mines data to, for example, see who a person talks to and how often, as well as the contents of communications. Lenddo's algorithm takes into consideration factors such as one-word subject lines (lack of attention to details), use of financial apps (taking finances seriously) and the proportion of selfies in a smartphone photo library (the use of a front-facing camera might indicate youth) (Bary, 2018). Lenddo also asks its members to provide a list of 'trusted friends' who are used as references. If a member's friends have not paid back loans, it affects the member's LenddoScore negatively. In April 2014, Lenddo teamed up with Scotiabank to give 100,000 Colombians access to credit cards based on LenddoScores.

Lenddo's typical loans are reported to be in the $300- $400 range. Lenddo, however, does not lend money itself. It charges other institutional lenders for assessing borrower creditworthiness. A consumer's LenddoScores and the local rates in that country determine the interest rates for a loan. The company claimed that compared to other resources, its loans are typically cheaper by about one-third (Beltran, 2011). Lenddo loans are reported to be used for education, medical and other purposes. Lenddo users are required to fill out an application, which takes about 15 minutes. Lenddo claimed that it responds as quickly as 24 hours (Beltran, 2011). In India, FICO formed a partnership with Lenddo to develop credit risk scores for consumers that have a limited or no formal credit history (FinTech Innovations, 2016). As of August 2015, Lenddo was issuing about 100,000 scores per month in India (Vageesh, 2016). India's peer-to-peer lending marketplace i-Lend also uses Lenddo's credit scoring model. As of September 2018, i-Lend had disbursed unsecured personal loans of about $12,000 (Ganguly, 2018).

Alibaba

Alibaba uses Big Data 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 transactions. Alibaba 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). As of October 2016, Alibaba had 423 million active buyers (Alizila, 2016).

In 2007, Alibaba launched AliLoan, by the China Construction Bank. The China Construction Bank had large amounts 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 bank so that the later could make better-informed lending decisions. The relationship terminated in 2011 when Alibaba reportedly asked the China Construction Bank to pay more for its credit information. Alibaba then used its own funds to lend via its AliFinance website. By mid-2012, AliFinance extended Chinese Yuan Renminbi 28 billion ($4.1 billion) in loans to more than 130,000 small businesses (Rabinovitch, 2013). During the three-year period following its creation, AliFinance issued over $16 billion in 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).

In June 2013 Alibaba launched a new fund management service called Yu'E Bao ("extra treasure" in Chinese) to compete with traditional bank 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 $1 billion in investments (economist.com, 2013). Yu'E Bao is an attractive alternative to traditional bank accounts, yielding annual interest rates of about 4.5 percent — significantly higher than the 0.35 percent rate on current deposits in banks — while enjoying the same liquidity (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 borrow directly from investors. Alibaba created a 14 billion Chinese yuan ($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, 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 programme 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 the Open Data Processing Service 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 Big Data 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 borrower 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, an Internet bank run entirely on the cloud. MYbank analyzes more than 3,000 variables and provided loans of $290 billion to about 16 million small companies by mid-2019 (aljazeera.com, 2019). The average loan size from MYbank to SMEs and individual businesses is reported to be $1,469 (Cheng, 2019).

In a 2016 ranking by KPMG and the investment firm H2 Ventures, Ant Financial topped the global ranking for the hundred best performing FinTech companies (Xinhua, 2016). Rural users are one of the key target groups of MYbank (*Asia Times*, 2015). The company aspires to provide credit to farmers for buying agricultural machines and tools (Kshetri, 2016b). Ant Financial plans to issue loans of up to $800,000 to small businesses and consumers (Kshetri, 2016b). MYbank's goal is to extend loans to 10 million SMEs in five years (Kshetri, 2016b). MYbank's data comes Ant Financial as well as from credit evaluation companies, such as Zhima Credit (Kshetri, 2016b), 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 of default on debts, late returns of rented cars and transactions on the Alipay online payment service (Kshetri, 2016b). Alibaba is exploiting its massive amount of data related to online consumption in the offline setting by accessing the relevant data through Alipay (Kshetri, 2016b).

In 2014, Alipay handled payments worth $800 billion (Kim, 2014). It uses payment data from Alipay to support the activities of Ant Financial (Creehan, 2016). Alipay credits can be used to buy consumer goods with offline retailers. In June 2015, Ant Financial announced a partnership with Walmart stores in which the latter would accept Alipay mobile payments service. The partnership would start with 25 stores in Shenzen and cover all 410 Walmart stores in China by the end of 2015 (Kshetri, 2016b). Alibaba's Big Data research on customer preferences, behavioural 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 (Kshetri, 2016b). The costs associated with approving a small business loan is reported to be 2 yuan ($0.32), compared to more than 2,000 yuan ($318) at a traditional bank (Zhang and Woo, 2018). Alibaba has also made use of artificial intelligence and machine learning. It uses deep-learning technology[2] to detect fraud (Perez and Soo, 2017).

---

2   Deep-learning uses algorithms, which teach computers to learn by examples and perform tasks based on classifying structured as well as unstructured data, such as images, sound and text.

3. Big Data as a Tool for Increasing Access to Financial Services for Disadvantaged Groups: Observations From the Selected Cases

Gartner has defined Big Data in terms of three Vs: volume, velocity and variety. The following discussion examines how the characteristics and dimensions of Big Data are relevant in the context of facilitating access to financial products for economically-active low-income families and micro-enterprises in emerging economies (Table 2).

Volume

A colossal increase in the digitization rate of Southern countries has occured, which has led to the availability of a large amount of data and information to assess the creditworthiness of individuals and enterprises. Of particular importance is the rapid diffusion of cellphones. According to the International Telecommunication Union, cellphone penetration rate in developing countries was 102.8 percent in 2018. For least developed countries, the proportion was 72.4 percent. Mobile phone-related data often provide high quality, valuable information because a mobile phone is often the only interactive technology for most low-income individuals. The case of Kilimo Salama indicates that low-cost sensors are another key source of data that has helped increase the access to financial products for small-holder farmers. One estimate suggested 14 billion sensors connecting devices were in the world in early 2015 (forbes.com, 2015).

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 access to financial services among the poor. For instance, Agrilife provides "near-real-time information" on farmer ability to pay for services (capacity.org, 2013). In another example, as noted above, Alibaba's MYbank's "310 model" approves a loan in one second (Chataing and Kushnir, 2018).

Variety

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 Cignifi can recognize patterns in consumer phone calls, text messages and data usage which are used to predict lifestyle and credit risk profile (bigdata-startups.com, 2013).

Table 2: *Relevance of Big Data 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 low-income families
Volume	Enormous amounts of data created and gathered from a wide range of sources	Mid-2014: Lenddo's network had about one million members. A typical Lenndo credit application had over 12,000 data points, which could be used to assess creditworthiness. Kreditech: processes over 8,000 data points in real-time.
Velocity	Time-sensitive data collected, stored, processed, analyzed and acted on quickly (importance of speed).	Agrilife: farmer's needs assessed by digital platforms instead of household surveys. Alibaba's MYbank: "310 model" approves loans in one second.
Variety	Data in multiple formats: structured and unstructured.	Structured data (e.g., no./amount of transactions), unstructured data (e.g., Cignifi: text messages).

4. Different Categories of Financial and Non-Financial Information

Fintech companies are utilizing a wide range of information from diverse sources to assess a borrower's creditworthiness. Potential sources of data and information about a borrower include telecommunication providers, utility companies, wholesale suppliers, retailers and government agencies. In addition, financial institutions might have data that was previously overlooked (Baer, Goland, and Schiff , 2013). E-commerce marketplaces and telecommunications are two industries with large amounts of digital customers with potentially useful information. Banks can provide lending services for the existing customers of e-commerce sites: consumers and SMEs. E-commerce players can gain easy access to financing for working capital. If consumers know financing is available, they can buy big-ticket items, such as computers, refrigerators and TVs (Barquin, 2016).

In attempting to explain the potential uses of Big Data in assessing creditworthiness of low-income 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, ability to repay and willingness to repay (Baer et al., 2013). Table 3 presents the functions of different categories of financial and non-financial information used by lending organizations. Some variables may reflect more than one underlying dimension. Categories of nonfinancial/non-payment information that can act as proxies for more than one dimension of creditworthiness are presented in Table 4. For instance, an individual's financial decision-making is a function of psychological variables, such as self-regulation, consideration of future consequences and basic financial knowledge (Kozup et al., 2008). Prior researchers have noted that consumers who consider their behaviours' future consequences are willing to focus on long-term financial security. To achieve that, these type of consumers may sacrifice the pleasure of spending in the short-term (Howlett et al., 2008). Such psychological variables may thus reflect ability and willingness to repay.

Fintech companies are utilizing a wide range of information from diverse sources to assess a borrower's creditworthiness. Potential sources of data and information about a borrower include telecommunication providers, utility companies, wholesale suppliers, retailers and government agencies.

4.1. Identity

Identity-related information helps to ensure that facts about a potential borrower is the same as what is provided or described by the borrower. Such information would help reduce potential fraud. An important issue in this regard is the source as well as the content of information. Information coming from a third party often has a higher degree of source credibility and trustworthiness compared to information that is self-reported by the potential borrower.

Identity-related information helps to ensure that facts about a potential borrower is the same as what is provided or described by the borrower. Such information would help reduce potential fraud.

Some encouraging technological developments are likely to further improve this landscape. A South African mobile payments provider was reported to be piloting the use of location data as a low-cost mechanism to validate self-reported addresses. To do so, it looks at a cellphone's nightly location patterns (Ehrbeck, 2015). Biometrics, such as eye scanning and fingerprinting, can be used to address the problem of producing identification for illiterate clients who are unable to provide a signature. Some Big Data companies are taking measures to create a unique 'financial identity' for farmers, which can be used by financial institutions in the process of assessing and quantifying credit risk. The Policy and Economic Research Council and Experian MicroAnalytics have developed a tool called Financial Identity Risk Management. Farmers can opt-in to provide biometric data, such as fingerprints, and authorize access to their mobile transactions and other data, such as utility bills and phone records. The Policy and Economic Research Council and Experian MicroAnalytics use the data to create a credit score using an algorithm. The credit score is passed to partner financial institutions (Babcock and Satham, 2014).

One proposal that was discussed in 2000 by the US Senate Committee on the Judiciary, Subcommittee on Technology, Terrorism and Government Information would have required the use of biometric technologies, such as fingerprints and voice patterns, to verify credit applicant identity (US Senate Committee on the Judiciary, 2000). Implementation of this requirement would require each lender to invest in biometric technologies and techniques. Analysts suggested at the time that the cost of such technologies would have been very high and might have even exceeded the total losses associated with identity theft (TowerGroup, 2002). Thus, not so long ago, the costs were likely to have exceeded the benefits. With the current availability of more effective solutions, financial institutions have greater incentive to adopt them and get rid of less efficient techniques.

4.2. Ability to Repay

To determine the ability to repay a loan, possession of the means needs to be assessed. Factors such as income and current debt load are some examples of proxies. Accion is reported to be working with companies whose data suggest that individuals living in rural areas who top up their cellphones on the same day every week and who pass by more than two mobile phone masts during the week are likely to have more reliable financial habits than those who top up irregularly and do not travel (Economist.com, 2014). Prior research in geography indicates that one reason why highly-mobile people move is related to the search for better economic opportunities. Such people often leave places that have unsatisfactory employment opportunities and move toward places where opportunities are better (Lonsdale and Archer, 1997).

Various proxies of social capital are also being used as a measure of user ability to repay. Some researchers have measured a person's social capital by counting the number of contacts to specific groups or individuals (Glaeser et al., 2001). The idea is that creating and maintaining contacts with people or groups entail investment of time and other resources (Becker, 1964). LendUp and Moven take into account the number of social media friends a borrower has as a proxy for the ability to pay (Jeffries, 2014).This simple counting, however, can be considered only a crude measure of social capital. To refine this measure, Big Data companies are relying on the quality of the contacts. Some use the reputation of social media friends to predict a user's ability to repay. The principal is that a borrower's likelihood to repay can be predicted by their friends' credit reputation (Cullerton, 2013). For instance, LendUp uses the frequency of social media interactions of a borrower as an indicator of stability (sjnewsonline.com, 2013).

4.3. Willingness to Repay

The willingness to pay is different from the ability to pay. For instance, such payments may disrupt households' usual consumption patterns or deplete their assets, putting the borrowers at risk of poverty (Russell, 1996). Possession of means thus may not be a sufficient condition to pay a loan. It is also important to assess whether the borrower has a strong disposition or inclination towards paying down debt. Past credit history and payment behaviour of a borrower are used to assess the willingness to repay.
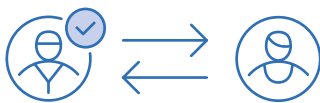
FinTech companies are launching services that are likely to help consumers address challenging behaviours and learn new behaviours so that they will be assessed favorably in terms of the willingness to repay.

FinTech companies are launching services that are likely to help consumers address challenging behaviours and learn new behaviours so that they will be assessed favorably in terms of the willingness to repay. One such example is RevolutionCredit, whose long-term vision is to promote financial inclusion in the developed as well emerging markets. RevolutionCredit describes its programme like 'traffic school' for credit. Customers can participate in enhanced education to compensate for a minor mistake. The company offers bite-sized, gamified financial education videos at the point of transaction to improve individual use of credit cards. Customers can improve their credits by watching a series of one-minute videos and passing corresponding tests (Faz and Noor, 2014). Accion, which as of 2015 helped create 63 microfinance institutions in 32 countries, invested $10 million in RevolutionCredit (Prnewswire.com, 2014).

Table 3: *Functions of different categories of financial and non-financial information.*

	Financial information	Non-financial information
Identity (to reduce fraud)	Third party information: high degree of trustworthiness	From government agencies (such as China's WeBank using data from the Ministry of Public Security). A South African mobile payments company: cellphone's nightly location patterns. Financial Identity Risk Management tool developed by the Policy and Economic Research Council and Experian MicroAnalytics: Farmers can opt-in to provide biometric, such as fingerprints.
Ability to repay	Income /current debt: Prepaid-phone minutes purchase patterns: steady/uneven cash flow (e.g., topping up cellphones the same day every week). Estimates of future production (collateral). Mobility: passing by more than two mobile phone masts during the week.	Reputation of friends. Hobbies (e.g., time spent playing video games).
Willingness to repay	Credit history: review of utility, rent, telephone, insurance and medical bill payments.	Items bought (e.g., buying diapers=responsibility). Customer reviews (for organizations). Books read (used by the Chinese government). RevolutionCredit: participation in enhanced education.

Table 4: *The roles of non-financial/non-payment information in accessing creditworthiness.*

Identity	Ability to repay	Willingness to repay
Biometric information (Financial Identity Risk Management tool developed by the Policy and Economic Research Council and Experian MicroAnalytics: Farmers can opt-in to provide biometric, such as fingerprints). Information from government agencies. Source: first party data. Cellphone's nightly location patterns. Remote sensing data, such as satellite imagery and photos.*	Meteorological monitoring processes, such as weather stations. Call detail record. Customer reviews (for businesses). Social media behaviours.** How people organize their lives: for example, phone contacts organized by both first and last names (Cheney, 2016). ** Attention to details: grammar and punctuation in a text message (Cheney, 2016).** Lenddo: one-word subject lines (lack of attention to details); use of financial apps (taking finances seriously). The proportion of selfies in smartphone photo library (the use of a front-facing camera might indicate youth) (Bary, 2018).** Items marked with an asterisk (*) in the 'identity' column.	Participation in financial education. Items bought (e.g., buying diapers=responsibility). InVenture's algorithms (e.g., users who wait to make calls when rates are lower (such as after 10 p.m.) are lower risk borrowers (Dwoskin, 2015). Kreditech: online gambling (credit risk); friends who have already repaid a loan (less risky). (Vasagar, 2016). Participation in financial education. Items marked with a double asterisk (**) in the 'ability to repay' column.

5. The Role of Big Data in Reducing Information Opacity and Transaction Costs

Table 5 presents the role of Big Data in reducing information opacity and transaction costs and compares the approaches of some Big Data firms discussed in the last section. As it is clear from Table 5, diverse sources of data are being used to reduce information opacity.

Some categories of information, especially nonfinancial/non-payment information, such as reputation of friends, hobbies and attention to detail (Tables 3 and 4) are of low or no marginal value in assessing potential borrowers' creditworthiness in industrialized countries, yet they play a significant role in the decision-making of FinTech companies serving the poor in Southern countries. This is due to the lack of availability of consumer financial/payment information in Southern countries. This nonfinancial/non-payment information is being used to reduce informational opacity, which has been a key barrier to access financial services for the poor (Stiglitz and Weiss, 1981).

Transaction costs are reduced by digitizing lending activities and/or minimizing or eliminating physical interaction between the lender and the borrower. For instance, the lack of an instantaneous credit scoring tool means that lenders need to go to people's homes and places of business and evaluate net worth and ability to repay. All these processes are unnecessary in the Big Data environment. For instance, First Access, which introduced its service in Tanzania, mines data stored on borrowers' mobile phones to recommend a loan amount to banks and microfinance lenders within 90 seconds, similar to Cignifi (Davis, 2016). First Access has developed a cloud-based platform that analyses how often a potential borrower replenishes airtime, buys data and interacts with social networks. The company charges about $1.25 per transaction to lenders. Using the information provided by First Access, lenders, on the average, are reported to save between $12 and $16 per evaluation (Browdie, 2013). Examples such as this demonstrate that the effects of data on reducing transaction costs are higher for borrowers of small amounts, who experience high borrowing costs measured as a proportion of the loan amount (Adams and Nehman, 1979). As prior researchers have pointed out, it is important to adopt innovative approaches in financial services to reduce transaction costs for such borrowers (Rogaly, 1996); Bid Data-based solutions are among such approaches.

The costs associated with collecting small cash payments act as a barrier for financial institutions to offer financial services to low-income and rural consumers. According to the Consultative Group to Assist the Poor, digitizing these transactions with mobile money accounts can reduce the cost of these payments to $0.10 per transaction (Faz and Noor, 2014). The rapid diffusion of cellphones and advances in cloud computing and data crunching technologies have lowered the cost of lending money to low-income people, as well as the costs of moving and storing money. Thanks to these innovations and advances a number of financial services, such as mobile wallets, crop insurance and new types of microloans, are available to low-income people.

The use of Big Data to assess creditworthiness is shaped by various institutional structures existing in Southern countries. For instance, the use of alternative data from non-financial firms may present competition and anti-trust issues in many countries. This is especially the case when such firms have strong market power, whereas national credit bureaus share credit scores with other financial institutions, non-financial firms are not required to do so. The development of alternative sources for scoring could thus encourage dominant players to create closed networks with data access limited to their own FinTech affiliates. In this way, the network effects required to stimulate SME finance are likely to be weakened (Creehan, 2016).

The rapid diffusion of cellphones and advances in cloud computing and data crunching technologies have lowered the cost of lending money to low-income people, as well as the costs of moving and storing money.

Table 5: *The role of Big Data in reducing information opacity and transaction costs: comparing some Big Data companies analyzed in this paper.*

	Reducing information opacity	Reducing transaction costs	Barriers to access to financial services overcome
Agrilife	The credit appraiser uses a range of data about a farmer, such as produce and status of farms, to assess creditworthiness. The input provider makes decisions on credit.	A farmer can make credit requests via a mobile phone.	By September 2013: Facilitated over $2 million in revolving credit lines to about 120,000 small farmers.
Kilimo Salama	Modernized solar power and computerized gauges/farmers linked to nearest weather station.	Transaction done via cell phones. No requirement to file a claim. No farm visits.	By 2014: covered close to 200,000 farmers in Kenya, Rwanda and Tanzania.
Cignifi	Looks at patterns in consumer phone calls, text messages and data usage to predict credit risk profile.	Works with mobile carriers and financial institutions in emerging markets.	Working with mobile carriers and financial institutions in several emerging markets to determine consumer credit risk based on cell phone usage data.
Monedo	Processes over 8,000 data points in real-time.	Does not require external credit bureau for identification, fraud detection and scoring decisions.	By 2015: served more than 800,000 customers and issued over two million loans to customers who lacked credit history to borrow money through other means.
Kueski	Partnership with Mambu Big Data team: machine learning methods using an algorithm. Alternative information.	Applying for a loan: computer/smart phone No personal meeting with the borrower.	As of 2019: provided more than 1.5 million loans worth $220 million.
Lenddo	Uses over 12,000 data points to assess borrower creditworthiness.	Charges institutional lenders for assessing borrower creditworthiness. Loans are typically cheaper by about one-third (Beltran, 2011).	2014: teamed up with Colombia's Scotiabank to give 100,000 Colombians access to credit cards based on LenddoScores. As of August 2015: issuing about 100,000 scores per month in India.
Alibaba	Utilizes its huge online ecosystem, with 37 million small businesses (Early 2015) and; 423 million active buyers (late 2016).	MYbank: Internet only bank. Open Data Processing Service: all decisions related to granting a loan made by the system without human intervention. Every loan is estimated to cost two yuan. Use of artificial intelligence to detect fraud.	Costs associated with approving a small business loan: $0.32. By mid-2012 AliFinance extended $4.1 billion in loans to over 130,000 small businesses.

6. The Role of South-South Cooperation in Leveraging the Impact of Big Data on Financial Inclusion

38% of global organizations are prepared to handle sophisticated cyber attacks.

As previously discussed, Big Data plays a crucial role in reducing information opacity and transaction costs for disadvantaged groups in the Global South, thus helping expand their access to finance. However, due to its inherent nature of big volume, Big Data could bring risks to data security. According to an Information Systems Audit and Control Association International report, only 38 percent of global organizations are prepared to handle sophisticated cyber attacks. Fintech company use of Big Data also could pose a risk of misusing personal data, discriminatory customer profiling systems, abusing new technology to support illicit transfers and other associated electronic fraud (IBRD/World Bank, 2018:18). Although more developing countries are investing in cyber security measures, they still fall behind in establishing good data governance frameworks in financial services.

South-South Cooperation could accelerate efforts to improve Southern countries' data governance, in addition to data security. In combatting financial cybercrime, for example, the Inter-American Committee against Terrorism was created to share information and expertise as a response to cyber incidents in the Americas. Moreover, South-South Cooperation is expected to contribute to protecting developing countries' data and establishing effective data governance policies that are not necessarily subject to developed country interests. For example, Free Basics, originally designed as a Facebook app to provide free Internet access in developing countries, is accused of harvesting a great amount of metadata about users and violating the principles of net neutrality (Guardian, 2017).

According to research by citizen media and activist group Global Voices, digital access for uses to western corporate content provided by large US companies was limited. There are also limitations in terms of language. In Ghana, for example, large US company content was only provided in English even though other languages, such as Twi and Hausa, are commonly spoken. In this regard, South-South cooperation could enable developing countries to play a more active role in governing user metadata, providing digital access tailored to local needs and in local languages and promoting an open Internet. In addition, with the increasing use of e-commerce, financial data stored in the cloud by big corporations, such as Amazon, is hard to govern. Big cloud companies have servers in many countries, but those countries differ in the extent of data governance regulations. Data could be stored in a developing country but might be processed in a developed

country whereas the former might not have adopted data security laws, but the latter could have already taken appropriate measures. Many legal questions remain unanswered, such as when data is regarded as exported (CIGI, 2019). Based on the agreed principles and objectives of South-South cooperation, such as equal partnership and active engagement of mutual parties, South-South cooperation has great potential to play a complementary role in establishing multilateral data governance regimes.

Furthermore, South-South cooperation could be leveraged to promote gender equality along with the use of Big Data to expand financial access. How Big Data could help disadvantaged groups, such as women, to have greater access to finance has already been discussed, but its benefits could reach even a greater number and wider groups of people through South-South cooperation. A strong positive correlation is already proven between gender equality and economic prosperity. Global Gross Domestic Product could increase by $12 trillion by 2025 if the gender gap in economic participation were reduced by only 25 percent (McKinsey, 2015). It would take reassured efforts of Southern countries to take greater advantage of Big Data in financial inclusion. For example, some middle-income Southern economies could share its valuable experience in achieving economic growth through women's economic empowerment. Pre-conditioning gender equality efforts as a key step to South-South cooperation could accelerate Big Data's effects in closing gender disparities in access to finance (Broussard, 2019).

South-South cooperation could also be a viable and powerful way to expedite access to credit for SMEs, especially in global supply chain. As opposed to traditional banking, Supply chain finance values the strength, longevity, performance and mutual dependence of relationships in a supply chain when assessing risks (Global Business Outlook, 2019). Supply chain finance helps change how large corporations are assessing suppliers and building sustainable partnerships between buyers and sellers. Through facilitation of knowledge sharing, partnerships and investments based on its key principle of mutual partnership, South-South cooperation could expedite supply chain finance for SMEs. Supply chain finance enlarges the impact of Big Data in evaluating the creditworthiness of SMEs as such source of finance helps financial providers set terms based completely on supplier or buyer historical performance without relying on credit ratings (Sreedhar, 2015). Some practices used with South-South cooperation in supply chain finance for SMEs are noticeable in China's Belt and Road countries. Singaporean financing platform SmartFunding provides digital invoice financing to directly fund SMEs without charging typical 3-7 percent rates for intermediary costs (Hartung, 2017).

$12 trillion

Global Gross Domestic Product could increase by $12 trillion by 2025 if the gender gap in economic participation were reduced by only 25 percent (McKinsey, 2015).

Implications

The preceding discussion points to an emerging trend in the use of Big Data to access the creditworthiness of disadvantaged groups and facilitate access to financial services. Some of the Big Data companies discussed in this paper are building their own databases (Agrilife, Alibaba, Kilimo Salama), while others rely on external data sources (Cignifi, Kueski, Lenddo, Monedo).

This section begins by reviewing the research questions addressed in this paper and findings so far (Table 6).

Table 6: *Revisiting the research questions*

Research question	Findings
RQ1 How can Big Data help reduce information opacity and transaction costs for disadvantaged groups in the global South?	Information from a variety of sources, such as individuals' phone calls, text messages, purchases and data usage, as well as from weather stations and satellites, are used. Machine learning methods are used. The process is completely automatized (e.g., applying for a loan via computer/smart phone). No personal meeting is necessary. Use of artificial intelligence to detect fraud (e.g., Alibaba, Kueski).
RQ2 How are different dimensions of Big Data related to disadvantaged group access to finance in the global South?	Various Big Data dimensions affect disadvantaged group access to finance in different ways. Some Southern countries, have a huge data pool (volume). Big Data can make it possible for disadvantaged groups to get loans quickly (velocity). . It is possible to use various types of unstructured data, for identification and to assess the ability to repay the loan (variety).
RQ3 What are the roles that formal and informal institutions play in facilitating the operations of Big Data companies in the financial and banking sectors of the global South?	Formal and informal institutions provide a conducive environment for the operations of Big Data companies in most Southern countries. For instance, Chinese consumers are less concerned about handing over personal data (Davis, 2016). Big Data companies face challenge in entering the industrialized world. The models used by the Big Data firms in assessing borrower creditworthiness and funding loans cannot be applied in the many Western countries (e.g., the Equal Opportunity Credit Act in the United States, the risk of discrimination against certain groups of individuals, etc.).
RQ4 How is South-South cooperation being leveraged to enlarge the impact of Big Data on financial inclusion?	South-South cooperation reaffirms the impact of Big Data in expanding financial accesses to disadvantaged groups, such as women and SMEs. It also could be leveraged to overcome the challenges of Big Data use in financial inclusion, such as data privacy and security issues. Based on its principles and objectives, South-South cooperation is expected to play a role in improving data governance and establishing data governance regimes that are not necessarily subject to developed country interests.

Regarding RQ1, the above discussion has made clear that Big Data can help reduce transaction costs and information opacity, which are key barriers to financial services for disadvantaged groups. Commenting on how Lenddo's models have addressed barriers to financial services in Southern countries, Arjuna Costa of Omidyar Network, Lenddo's investor noted: "By using non-traditional data to make thousands of creditworthy consumers discoverable to lenders for the first time, Lenddo is revolutionizing the way we reach, assess and ultimately extend credit to consumers with no formal credit histories in developing markets" (omidyar.com, 2015).

Especially companies in China have databases bigger than any other country. For instance, Tencent's WeChat had over 1.08 billion active users monthly in 2018Q3 (Lee, 2018). Both Big Data companies and financial institutions in China are taking initiatives to be in cell [I] in Table 1. That is, Big Data organizations want to provide financial services by themselves. Financial services providers, on the other hand, are developing big data capabilities. In this way, they are invading each other's territories. For instance, e-commerce companies, such as Alibaba, have launched financial products and services, such as money market funds and consumer loans. As discussed above, the Big Data companies discussed in this paper, however, are either building their own databases or relying on data from other sources.

In China, traditional banks are diversifying businesses by investing in e-commerce and transforming themselves as Big Data companies. For instance, as of early 2012, the Chinese financial industry was estimated to have more than 100 terabytes of structured and unstructured data (IDC, 2012). As of March 2014, the Industrial and Commercial Bank of China, the country's largest lender, was reported to have over 4.9 petabytes of data. Likewise, the Agricultural Bank of China was estimated to generate 100 terabytes of structured data and 1 petabyte of unstructured data in 2014 (ABC, 2014). Similarly, in 2014, the Bank of Communications reportedly handled about 600 gigabytes of data daily and had a storage capacity of more than 70 terabytes (BOCOM, 2014). According to Forbes' list of the World's Biggest Public Companies (www.forbes.com/global2000/list/), the Industrial and Commercial Bank of China ranked #1 and the Agricultural Bank of China ranked #3. Two other Chinese banks, the China Construction Bank and the Bank of China, ranked #2 and #4, respectively. Most traditional banks, however, in other Southern countries are significantly smaller than those in China and thus lack the capability to incorporate Big Data-based decisions. Likewise, Big Data players outside China lack resources and capabilities and are unprepared to invade the businesses of traditional banks.

As noted above, every loan is estimated to cost 2 yuan for Alibaba, which is significantly lower than the $1.25 per transaction charged to lenders by First Access. While Tanzanian banks are reported to save $12-16 on average in assessing a potential borrower's creditworthiness thanks to services provided by First Access, the costs are still prohibitively high for the small-scale loans that low-income people typically can afford. For instance, consider the microfinance model, Village Savings and Loans Associations, which had reached 2.3 million people in 26 African countries by 2011. The average loan size of a Tanzania Village Savings and Loans Association member was $61 in Tanzania and $37 in Uganda (Hendricks and Chidiac, 2011). Moreover, while China's Internet-only banks, such as the Alibaba Group's MYbank and Tencent's WeBank, can greatly reduce the cost of disbursing and monitoring small loans. Such arrangements do not exist in most other countries.

A further mechanism that has contributed to the reduction of transaction costs is the use of artificial intelligence and machine learning to detect fraud. Stewart Langdon, Partner at LeapFrog Investments, which has invested in Cignifi, commented: "Up until now, the world's low-income consumers have had a very hard time borrowing money because they don't have credit histories. Cignifi addresses this through a machine learning platform that integrates with the world's most accessible infrastructure – mobile networks – to give people a digital identity. This financial technology has the potential to revolutionize how we assess the creditworthiness of billions of people globally and drive down customer acquisition costs" (cignifi.com, 2017).

A further mechanism that has contributed to the reduction of transaction costs is the use of artificial intelligence and machine learning to detect fraud.

Likewise, as mentioned, Alibaba uses deep-learning technology to detect fraud (Perez and Soo, 2017). The company's losses related to fraud are one in 1 million (Perez and Soo, 2017).Kueski's artificial intelligence algorithms identify unusual patterns of potential borrowers that have applied for a loan. Data and information submitted by users to its platform and those gathered from other sources are used for this purpose (López, 2020). As to RQ2, various Big Data dimensions are uniquely related to disadvantaged groups' access to finance. Some Southern countries, such as China, have a data pool that is bigger and deeper than other countries in the world (volume).

Big Data can speed the process by which disadvantaged groups can get loans. For instance, commenting on the roles of these technologies in overcoming disadvantage group access to finance, Kueski's founder and chief executive Adalberto Flores put the issue this way: "We believe that customers should be able to access credit, quickly and easily, when they need it, without having to take time out of work to visit a branch, or wait days for the outcome of an outdated credit review process" (prnewswire.com, 2014). In the future, in addition to structured data (variety), it is possible to use unstructured data, such as a satellite image of the roof of a farmer's house or the size of a plot of agricultural land, can be used for identification and to assess the ability to repay the loan (variety).

Regarding RQ3, many Southern countries have institutional foundations that have supported the growth of Big Data 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 GlobalTimes noted that without the firewall "China would become therealm of Google China, Yahoo China and Facebook China" (wsj.com, 2015). A related point is that compliance with government 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 a monopoly of firms, such as Alibaba and Tencent. Thee firms that are endowed with a monopoly position in the market are in a position to attract a large number of customers on their websites and collect a vast array of information about them. Many other Southern countries lack such conditions.

Internet-only banks, such as MYbank and WeBank, face barriers to expand their operations. They are required to serve other banks' customers because they lacked physical outlets and thus can not verify identities of every new customer. While the banks claim they can verify new customers remotely through facial recognition technologies, regulators have been adamant that new customer identities need to be verified in person (Hongyuran, 2015). The People's Bank of China rejected requests from Internet-only banks for remote verification. The Internet-only banks needed to work with traditional banks by linking accounts (Hongyuran, 2015).

Big Data companies, such as Lenddo, may face challenges in entering the industrialized world. For instance, the United States has strict credit underwriting regulations, which means that the models used by the above Big Data firms in assessing borrower credit worthiness and funding loans cannot be applied in the US. To use a credit scoring system, Big Data companies must prove the efficacy of their risk models. They also need to adhere to regulations related to information gathering and distribution

and make sure that they protect consumer privacy. The Equal Opportunity Credit Act presents a challenge of another magnitude. This statute prohibits lending companies from inadvertently discriminating against individuals based on age, sex, race or religion. The risk of discrimination against certain groups of individuals is likely to be high in models that use scores based on data from social networks (Woodruff, 2015). FICO has no plan to use non-payment data for US consumers (Andriotis and Demos, 2016).

Big Data deployments in the global South vary widely in terms of project capital, resource intensiveness, sophistication, complexity, performance and impact. Structural differences between the Big Data industries in China and other Southern countries are worth mentioning. To illustrate this point, compare the Big Data deployments by China's Alibaba and Kenya's Agrilife. As of 2015, Alibaba had a market value of about $233 billion, making it the world's third largest global public Internet company only behind Apple and Google (Schwarzmann, 2015). Like other Chinese Internet companies, Alibaba has access to huge amounts of data, providing a strong foundation to develop Big Data-based financial services. It is fair to say that among the developing world-based firms, Alibaba's Big Data tools are among the most advanced and sophisticated. In contrast, Agrilife is technically less sophisticated with a simpler and cheaper platform. Data volumes handled by Agrilife are not as big as Alibaba and actions are taken on a near real-time basis rather than a real-time manner. For instance, in 2014, Agrilife served about 120,000 smallholder farmers (Kshetri 2014) compared to 314 million customers served by Alibaba (Smith, 2020).

Especially impressive is African big data firms discussed above, which have created innovative solutions using simple technologies to help disadvantaged groups. In this regard, it is worth mentioning , Foreign Policy magazine's survey with the world's top Internet experts on Internet-related innovations. Seven percent of the experts viewed Africa as "the most innovative place for Internet-related technology." The corresponding proportions for other regions and economies were: Europe: 4 percent; China: 4 percent; India: 7 percent; and the Pacific Rim: 5 percent. The experts viewed Africa's Internet-related innovations as: "On-the-ground solutions designed by communities for communities" (Foreign Policy, 2011). Earlier discussion about big data solutions launched by Africa-based Big Data firms, such as Agrilife and Kilimo Salama, supports this observation.

While cellphones are mainly viewed as a communication tool, they can potentially play an even more important role in Southern countries, especially for low-income people, by providing information that can increase access to finance. Given the limited success, to date, of other measures to increase low-income people's access to financial services (Azmeh, 2018; IFC, 2019; Miranda and Gonzalez-Vega, 2011), one way to address this situation would thus be to take measures to accelerate the diffusion of digital technologies, such as cellphones, among low-income households. Equally important is to attract foreign FinTech companies that provide such services and to develop local capacity on this front.

A complaint is that Big Data analytics used to predict potential borrower creditworthiness rely on correlations rather than causations. It is not easy to determine which correlations shown by Big Data tools are random and which ones may reflect responsible financial and consumption decisions (Zetsche et al., 2018). In this regard, it is reasonable to expect that models based on Big Data tools will be more refined in the

future so that only variables that are related to responsible financial decisions will be used to assess creditworthiness. Policymakers should work with Big Data companies to refine algorithms so that they can accurately capture responsible financial decisions.

While various FinTech companies have been successful in some Southern countries, their models may not be transferable to other developing countries. For instance, Alibaba can provide loans at significantly lower costs because it uses its own technologies instead of expensive foreign software and has no physical branches (Zhang and Woo, 2018). FinTech companies in most developing countries may lack capabilities to do this.

Regarding RQ4, South-South cooperation has been and could be further leveraged to complement the limitations as well as accelerate benefits of Big Data's use in financial inclusion. Big Data brings about inherent challenges, such as vulnerability to data breaches. Many global corporations who store and collect big data in developing countries are also susceptible to work on developed country terms, as was seen with Facebook's Free Basics app. South-South cooperation has contributed to strengthening financial inclusion policy through better management and governance of data. For example, the Alliance for Financial Inclusion, a network of policymakers and regulators from over 80 countries in the global South, has created a working group on "data and measurement that is exploring how countries can better gather financial inclusion data, and more specifically demand-side and usage information" (Gardeva, 2012). Such sharing of knowledge and expertise among Southern countries has supported the development of each nation's policy, regulation and strategies for financial inclusion.

South-South cooperation could contribute to fueling Big Data's role in closing gender gaps in financial access. It is recommended that some leading economies in Southern countries, such as China, publicize their gender equality efforts in utilizing Big Data to expand financial access to disadvantaged groups. Publicized proof of correlation between gender parity and economic prosperity would lead other Southern countries to invest in and benefit from Big Data's role in financial inclusion. The expected benefits of gender parity in financial access could be reassured through pronounced gender equality efforts from leading global South financial service providers.

South-South cooperation could also advance the impact of Big Data's role in financing for SMEs. Fintech companies from Southern countries develop and provide new products tailored to the needs of SMEs while fully leveraging the role of Big Data in reducing information opacity and transaction costs. Accordingly, they became major new players providing supply chains for finance for SMEs in other Southern countries. A key point here is that firms based in Southern countries such as China are in a position to reconfigure their resources to operate in other Southern countries. They can easily adapt the business models used in their domestic markets China's Belt and Road Initiative could be a good example where Southern countries along the Belt and Road Initiative are building a collaboration platform for digital financial inclusion. Leading Chinese digital financial inclusion organizations' partnership with other Southern countries is highly desirable in that it could provide greater financial access to SMEs, thus help their incorporation into the global supply chain, as well.

7. Conclusion

The present study extends as well as complements prior research that examined the role of Big Data in increasing access to financial services to disadvantaged groups in the global South (e.g., Kshetri, 2016b). The above discussion indicates that one of the main reasons why low-income families and micro-enterprises in 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 this financially disadvantaged group. Thanks to Big Data, the situation is improving, but there much room remains for improvement. For instance, a comparison of Tanzanian banks with China's Alibaba indicates that the former group exhibits higher transaction costs than the latter. This means that the full potential of Information and Communications Technology has not yet been realized, such as the increased need for the development of software and algorithms to model relationships between a user's online behaviour and their creditworthiness.

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 Big Data can help financial institutions overcome two major challenges discussed above: reducing information opacity and transaction costs. Thanks to Big Data, farmers and agri-businesses have access to affordable and attractive financial services. Overall Big Data has potential to improve disadvantaged group access to credit, in terms of both quantities and terms.

Observers have noted that many practitioners lack skills and knowledge in various Big Data areas and they are not equipped to gain meaningful insights from diverse datasets. This situation is even more pronounced in the context of Southern economies. 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 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 citizen identity and other information with private companies (Baer et al., 2013).

Financial institutions that have a huge amount of information about individuals and businesses and/or those that possess Big Data analytic skills, that develop efficient and effective algorithms for assessing creditworthiness and that can reduce fraud are likely to benefit from this trend. The business models and approaches discussed above have widened the availability of financial services to economically-disadvantaged groups but have not yet been able to do so for the poorest of the poor. Big Data firms still lack reasonable amounts of information on people with the lowest incomes, who do not yet own a cellphone or have a social networking account. There are also different forms of digital divides that are emerging. For instance, Kilimo Salama's weather monitoring stations are primarily installed in heavily farmed areas in Kenya and Rwanda. This means that areas that are not heavily farmed lack required data. To close the gap between those disadvantaged groups and others, the role of South-South cooperation is more important than ever. Facilitation of knowledge sharing, partnerships and investment among Southern countries could complement Big Data's limitations while reassuring its positive impact on financial inclusion.

References

Aaronson S. (2019). "Data is a Development Issue." CIGI Paper Series No. 223. www.cigionline.org/publications/data-development-issue; accessed 23 March 2020.

Agricultural Bank of China (ABC). (2014). "Annual Report of ABC 2013." Beijing. www.abchina.com/en/investor-relations/performance-reports/annual-reports; accessed 18 November 2019.

Adams, D. W., and Nehman, I, G. (1979). "Borrowing Costs and the Demand for Rural Credit." *The Journal of Development Studies*, 15(2), pp. 165-176.

africanbusinessreview.co.za 2020. Safe Farming: An innovation in technology, February 18. www.africanbusinessreview.co.za/technology/software/safe-farming-an-innovation-in-technology.

Alibabagroup.com. (2015). "Ant Financial Unveils China's First Credit-Scoring System Using Online Data." www.alibabagroup.com/en/news/article?news=p150128; accessed 18 December 2019.

Alizila. (2016, October 17). "Five Reasons Why Alibaba is a Data (Not E-commerce) Company." www.alizila.com/five-reasons-why-alibaba-is-a-data-company/; accessed 23 November 2019.

Aljazeera.com. 2019. "Jack Ma's $290bn Loan Machine is Changing Chinese Banking," July 29. www.aljazeera.com/ajimpact/jack-mas-290bn-loan-machine-changing chinese-banking 190729023517726.html; accessed 17 December 2019.

Alois, JD. (2020)0 Name Change: Kreditech Rebrands to Monedo, March 10, www.crowdfundinsider.com/2020/03/158551-name-change-kreditech-rebrands-to-monedo/.

Analytix, G. (2013). "Credit Information Sharing: Unlocking Access to Affordable Credit: Alternative Data to Drive Financial Inclusion, the 2nd Regional Credit Information Sharing Conference." Nairobi: Kenya; Miranda, M. J., Vedenov, D. V., 2001. *Innovations in Agricultural and Natural Disaster Insurance*.

Andriotis, A. and Demos, T. (2016, October 24). "Cellphone Use to Determine FICO Scores Abroad." *The Wall Street Journal*. www.wsj.com/articles/fico-to-let-some-credit-scores-reflect-nonfinancial-information-1477323342; accessed 11 December 2019.

Ansari, S. Garud, R. and Kumaraswam, A. (2016). "The Disruptor's Dilemma: TiVo and the U.S. Television Ecosystem." *Strategic Management Journal*. 37 (9), 1829-1853, http://onlinelibrary.wiley.com/doi/10.1002/smj.2442/abstract; accessed 3 December 2019.

Asia Times. (2015). "Jack Ma's Online Bank lifts Off," June 25. http://atimes.com/2015/06/jack-mas-online-bank-lifts-off/

AWS Case study: Lenddo. https://aws.amazon.com/solutions/case-studies/lenddo/.

Azmeh, C. (2018). "Foreign Bank Entry and Financial Development: New Evidence on the Cherry Picking and Foreign Bank's Informational Disadvantage Phenomena in the MENA Countries." *Cogent Economics and Finance*, Vol 6. Issue 1. www.tandfonline.com/doi/full/10.1080/23322039.2018.1452343; accessed 21 January 2020.

Babcock, H, L., and Statham, C. (2014). "The Agricultural Mobile Finance Revolution." Issue no. 76. http://ictupdate.cta.int/Feature-Articles/The-agricultural-mobile-finance-revolution/%2876%29/1392201374; accessed 10 December 2019.

Baer, T. et. al (2013). "New Credit-Risk Models for the Unbanked, "April 2013. www.mckinsey.com/insights/risk_management/new_credit-risk_models_for_the_unbanked; accessed 3 December 2019.

Bansal P, Corley K. (2012). "Part 7: What's Different About Qualitative Research?" *Academy of Management Journal,* 55(3), 509–513.

Barquin, S. and HV, V. (2016). "Building a Digital-Banking Business." April.

Bary, E. (2018). "How artificial intelligence could replace credit scores and reshape how we get loans," October 29. www.marketwatch.com/story/ai-based-credit-scores-will-soon-give-one-billion-people-access-to-banking-services-2018-10-09

Beck, T et. al. (2004). "Bank Competition and Access to Finance." *International Evidence Journal of Money, Credit, and Banking*, 36(3), pp. 627-648.

Becker, G.S. (1964). "Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education." *Columbia University Press*, New York, USA.

Beltran, L. (2011). "Lenddo Provides Loans for Consumers in Emerging Markets." July 25. www.pehub.com/2011/07/lenddo-provides-loans-for-consumers-in-emerging-markets/; accessed 2 December 2019.

Bank of Communications (BOCOM). (2014). "Annual Report of BOCOM 2013." Shanghai. www.bankcomm.com/BankCommSite/en/invest_relation/company_develop.jsp?type=report; accessed 22 December 2019.

Borst, N. and Li, C. (2016). "Can Private Capital Transform Banking in China?" *Federal Reserve Bank of San Francisco*. www.frbsf.org/banking/asia-program/pacific-exchange-blog/china-private-banks-transform/; accessed 13 December 2019.

Broussard, J. (2019). "China's Potential to Promote Gender Equality through South-South Cooperation" Georgetown Journal of International Affairs. www.georgetownjournalofinternationalaffairs.org/online-edition/2019/3/28/chinas-potential-for-promoting-gender-equality-through-south-south-cooperation; accessed 20 March 2020.

Browdie, B. (2013). "First Access, Cignifi Use Big Data to Bring Economic Security to the Unbanked." www.americanbanker.com/issues/178_246/first-access-cignifi-use-big-data-to-bring-economic-security-to-unbanked-1064537-1.html; accessed 10 January 2020.

Capacity.org. (2013). "Agrilife Platform." www.capacity.org/capacity/opencms/en/topics/gender-and-social-inclusion/agrilife-web-platform-kenya.html; accessed 19 January 2019.

Chaia, A., et. al. (2010). "Counting the World's Unbanked." *McKinsey Quarterly* 2, pp. 98-99.

Chen, Y. (2014). "Alibaba Arm to Create $163 Billion Loans Marketplace." September 23. www.bloomberg.com/news/articles/2014-09-23/alibaba-arm-aims-to-create-163-billion-loans-marketplace; accessed 2 December 2019.

Cheng, E. (2019). "China Wants to Boost Loans to Small Businesses: Tech Companies May be the Answer." January 28. www.cnbc.com/2019/01/29/chinese-fintech-companies-find-new-opportunities-in-business-loans.html; accessed 10 January 2020.

Cheney, C. (2016). "How Alternative Credit Scoring is Transforming Lending in the Developing World." Deveximpact. September 8. www.devex.com/news/how-alternative-credit-scoring-is-transforming-lending-in-the-developing-world-88487; accessed 28 December 2019.

Chng ,J., Zhang, J. (2014). "Big Data: The Next Leading Edge in the Financial Industry." www.atkearney.com.tr/financial-institutions/ideas-insights/article/-/asset_publisher/LCcgOeS4t85g/content/big-data-the-next-leading-edge-in-the-financial-industry/10192?_101_INSTANCE_LCcgOeS4t85g_redirect=%2Ffinancial-institutions%2Fideas-insights%2F-%2Fasset_publisher%2FSxO47XYyZBF5%2Fcontent%2Fthe-new-reality-of-gcc-banking%2F10192; accessed 5 December 2019.

cignifi.com, 2017. Leapfrog and Omidyar Network target unbanked billions in Cignifi deal, June 23. www.cignifi.com/post/leapfrog-and-omidyar-network-target-unbanked-billions-in-cignifi-deal.

Costa, A., et. al. (2015). "Big data, Small Credit." *Omidyar Network*. www.omidyar.com/spotlight/what-big-data-small-credit; accessed 8 December 2019.

Creehan, S. (2016). "Can Fintech Fill Asia's SME Lending Gap? Federal Reserve Bank of San Francisco." October 16. www.frbsf.org/banking/asia-program/pacific-exchange-blog/can-fintech-fill-asias-sme-lending-gap/.

Cta.int. (2014). "Agrilife: Bringing the Agriculture Value Chain Finance Ecosystem to Life." July 14. http://blogs.cta.int/2014/07/14/agrilife-bringing-agriculture-value-chain-finance-ecosystem-life/; accessed 18 December 2019.
Cull, R. and Martinez, P. (2007). "Foreign Bank Participation and Crises in Developing Countries." *World Bank Policy Research Working Paper Series*, WPS4128.

Cullerton, N. (2013). "Behavioral Credit Scoring." *Georgetown Law Journal*, 101, 807-838.

Davis, T. (2016, June 20). "In China, Consumers Expose Private Data -- Including Nude Selfies -- for Better Loans." *The Dallas Morning News*. www.dallasnews.com/business/money/2016/06/20/in-china-consumers-expose-private-data----including-nude-selfies----for-better-loans; accessed 18 January 2019.

Delgado, R. (2016). "How Big Data Will Transform the Lending Industry." *Data Informed*. September 29. http://data-informed.com/how-big-data-will-transform-the-lending-industry/. Accessed on 13 January 2019.

Dovi, E. 2011."Boosting Domestic Savings in Africa," *Africa Renewal*, vol. 22, no. 3. www.un.org/ecosocdev/geninfo/afrec; accessed 28 January 2020.

Dwoskin, E. (2015). "Lending Startups Look at Borrowers' Phone Usage to Assess Creditworthiness." *The Wall Street Journal*, November 30. www.wsj.com/articles/lending-startups-look-at-borrowers-phone-usage-to-assess-creditworthiness-1448933308; accessed 4 January 2019.

Economist.com. (2013). "Giant Reality-Check Four of the world's Biggest Lenders Must Face Some Nasty Truths." www.economist.com/news/finance-and-economics/21584331-four-worlds-biggest-lenders-must-face-some-nasty-truths-giant-reality-check; accessed 19 December 2019.

Economist.com. "A Pionful of Dollars: The World's Poor Need the Stability and Security that Banks have Traditionally Offered, but Increasingly They Do Not Need Banks to Provide it." 15 November 2014. www.economist.com/news/briefing/21632441-worlds-poor-need-stability-and-security-banks-have-traditionally-offered; accessed 13 December 2019.

Economist.com. "A Nascent Tech Hub May Succeed by Solving Local Problems." 4 April 2015. www.economist.com/news/business/21647624-nascent-tech-hub-may-succeed-solving-local-problems-techs-mex; accessed 18 November 2019.

Ehrbeck, T. (2015). "Driving Financial Inclusion at 4G Speed." www.huffingtonpost.com/tilman-ehrbeck/smartphones-driving-finan_b_7011282.html; accessed 17 December 2019.

Eisenhardt, K. M. (1989). "Building Theories from Case Study Research." *The Academy of Management Review*, 14, 532–550.

Eisenhardt, K.M. and Graebner, M.E. 2007. "Theory Building from Cases: Opportunities and Challenges." *Academy of Management Journal* 50(1):25-32.

Engen, J. (2009). "Is Microfinance Ready for Its Next Big Leap*?"* *U.S. Banker*, 119(2), p. 19.

Eppler, M. J. (2006). "Managing Information Quality: Increasing the Value of Information in Knowledge-Intensive Products and Processes." *Springer*. Berlin.

Faz, X., and Noor, W. (2014). "New CGAP Research Points to the Potential Game-Changing Power of `Second Generation' Digital Payments Products. October 23. http://paymentweek.com/2014-10-22-new-cgap-research-second-generation-digital-payments-products-5923/; accessed 12 September 2019.

Financial Times. (2017). "FinTechs Could take 30% of Mexican Banking Market." www.ft.com/content/5e092fb8-3bb9-11e7-ac89-b01cc67cfeec; accessed 18 November 2019.

FinTech Innovations. (2016). "FICO and Lenddo Join Hands to Extend Credit Reach in India." www.enterpriseinnovation.net/article/fico-and-lenddo-join-hands-extend-credit-reach-india-1027757993; accessed 13 October 2019.

Flores, A. (2016, April 21). "Mexican Fintech Startup Kueski Raises US$35m." *LAVCA*. http://lavca.org/2016/04/21/mexican-fintech-startup-kueski-raises-us35m/; accessed 27 November 2019.

Forbes.com. (2015). "Why the Internet of Things Heralds the Next Great Economic Disruption." www.forbes.com/sites/joemckendrick/2015/03/25/is-the-internet-of-things-heralding-the-next-great-economic-shift/; accessed 10 October 2020.

Foreign Policy. (2011). "The FP Survey: The Internet." September/October (188):1–9.

Ganguly. P. (2018), P2P lending network i-Lend partners with Singapore-based Lenddo for credit scores October 03. https://economictimes.indiatimes.com/industry/banking/finance/p2p-lending-network-i-lend-partners-with-singapore-based-lenddo-for-credit-scores/articleshow/50806520.cms.

Gardeba, A. (2012). "Four Ways Big Data Will Impact Financial Inclusion." Center for Financial Inclusion 2020 Blog Series. www.centerforfinancialinclusion.org/four-ways-big-data-will-impact-financial-inclusion; accessed 20 March 2020.

Gartner.com. (2013). "*Big Data*." www.gartner.com/it-glossary/big-data/; accessed 18 October 2019.

Glaeser, E.L., et.al. (2001). " The Economic Approach to Social Capital." Discussion paper no. 1916. *Harvard Institute of Economic Research*, Cambridge, MA, USA.

Glaser, B. G. and Strauss, A. L. (1967). "The Discovery of Grounded Theory: Strategies for Qualitative Research."Aldine, Chicago.

Global Business Outlook (2019). "Supply Chain Finance – Good for Banks and SMEs." www.globalbusinessoutlook.com/supply-chain-finance-good-for-banks-and-smes/; accessed 25 March 2020.

Golder, P. N. (2000). "Historical Method in Marketing Research with New Evidence on Long-Term Market Share Stability." *Journal of Marketing Research*, 37 (2):156-72.

Gottschalk, L. (1969). "Understanding History: A Primer of Historical Method." New York: *Alfred A. Knopf*.
Government Information (2000). "ID Theft: When Bad Things Happen to Your Good Name. Hearing." www.gpo.gov/fdsys/pkg/CHRG-106shrg69821/html/CHRG-106shrg69821.htm; accessed 14 August 2019.

Guardian.com. (2017). "'It's digital colonialism': how Facebook's free internet service has failed its users." www.theguardian.com/technology/2017/jul/27/facebook-free-basics-developing-markets; Accessed 23 March 2020.

Greatrex H., et.al (2015). "Scaling up Index Insurance for Smallholder Farmers: Recent Evidence and Insights." CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), Report No. 14.

Greene, D. and David J.L. (1984). "A Research Design for Generalizing from Multiple Case Studies, Evaluation and Program planning." 7: 73-84.

Grossman, J. and Tarazi, M. (2014). "Serving Smallholder Farmers: Recent Developments in Digital Finance." Focus Note 94. Washington, DC: CGAP, June. www.cgap.org/sites/default/files/Focus-Note-Serving-Smallholder-Farmers-Jun-2014.pdf; accessed 12 January 2020.

H2 Ventures and KPMG (2019). "2019 Fintech 100: Leading Global Fintech Innovators." New York: *H2 Ventures and KMG*. https://h2.vc/wp-content/uploads/2019/12/2019Fintech100.pdf; accessed 25 January 2020.

Hamilton, D., (2016). "Canadian Banks work with (Not Against). FinTech Startups to Bring Innovation to the Industry: Report." March 11. www.itbusiness.ca/news/canadian-banks-work-with-not-against-fintech-startups-to-bring-innovation-to-the-industry-report/66625; accessed 19 November 2019.

Hartung, R. (2017). "Asia's Supply Chain Financing: From Tech Innovation to China's Belt and Road." March 13. www.theglobaltreasurer.com/2017/03/13/asias-supply-chain-financing-from-tech-innovation-to-chinas-belt-and-road/; accessed 25 March 2020.

Hendricks, L and Sybil, C. (2011). "Village Savings and Loans: A Pathway to Financial Inclusion for Africa's Poorest Households." *2011 Global Microcredit Summit*, Commissioned Workshop Paper, November 14-17, Valladolid, Spain.

HIQA: International Review of Data Quality Health Information and Quality Authority (HIQA). (2011). www.hiqa.ie/press-release/2011-04-28-international-review-data-quality; Accessed 09.06.16.

Holley, E. (2014). "Financial Services Looking Up, down Mexico Way." www.bankingtech.com/206102/financial-services-looking-up-down-mexico-way/; accessed 24 November 2019.

Hongyuran, W. (2015). "Closer Look: Baidu Joins Other Internet Giants to Open a Bank." November 12. http://english.caixin.com/2015-11-20/100876718.html; accessed 23 November 2019.

Hongyuran, W., 2015. "Internet-Only Banks in China Not Clicking Yet." October 8. http://english.caixin.com/2015-10-08/100860688.html; accessed 17 December 2019.

Howlett, E., Kees, J. and Kemp, E. (2008). "The Role of Self-Regulation, Future Orientation, and Financial Knowledge in Long-Term Financial Decisions." *Journal of Consumer Affairs*, 42(2), pp. 223–42. www.mckinsey.com/industries/financial-services/our-insights/Building-a-digital-banking-business?cid=other-eml-alt-mip-mck-oth-1304; accessed 11 December 2019.

HSBC (2018). "BRI: Rising to the Fintech Challenge." August 8. www.business.hsbc.com/belt-and-road/bri-rising-to-the-fintech-challenge; accessed 25 March 2020.
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; accessed 4 January 2020.

IFC. (2019). "Banking on SMEs: Trends and Challenges. Washington DC: International Finance Corporation." www.ifc.org/wps/wcm/connect/dd06b824-c38b-4933-9108-0c834f182fee/IFC+on+Banking+SMEs+Publication+June+2019.pdf?MOD=AJPERES&CVID=mSdrGtA; accessed 19 January 2020.

International Bank for Reconstruction and Development/ World Bank (IBRD/ World Bank). (2018) Global Financial Development Report 2017/2018: Bankers without Borders. Washington DC: World Bank Group. https://openknowledge.worldbank.org/bitstream/handle/10986/28482/9781464811487.pdf

Jack, W. and T. Suri. (2010). "The Economics of M-PESA." www.mit.edu/~tavneetM-PESA.pdf; Accessed on 13 October 2019.

Jappelli, T., and Marco P. (2002). "Information Sharing in Credit Markets: International Evidence." *Journal of Banking and Finance* 26, pp. 2023–2054.

Jayawardene, V., Sadiq, S. and Indulska, M. (2013). "An Analysis of Data Quality Dimensions." ITEE Technical Report 2013- 01. *School of Information Technology and Electrical Engineering*. The University of Queensland.

Jing, M. (2014). Alibaba, Lenders Team up for SME Financing. China Daily, July 23. www.chinadaily.com.cn/business/2014-07/23/content_17901230.htm; accessed 15 October 2019.

Kalan, J. (2013). "Tech Fix for Africa's Big Farming Challenge." www.bbc.com/future/story/20130408-tech-taps-africas-farm-potential; accessed 3 October 2019.

Kim, E. (2014). "A Company Spun Out of Alibaba is Now Worth as Much as $50 billion." *BusinessInsider.com*. www.businessinsider.com/ant-financial-closes-deal-that-gives-50-billion-valuation-2015-7; accessed 29 October 2019.

Kokalitcheva, K. (2014). "Cignifi Raises Funding to Model Credit Scores for Emerging Market Consumers." http://venturebeat.com/2014/09/05/cignifi-raises-funding-from-to-model-credit-scores-for-emerging-market-consumers/; accessed 27 October 2019.

Kozup, J, et. al. *(*2008). "The Effects of Summary Information on Consumer Perceptions of Mutual Fund Characteristics." *Journal of Consumer Affairs, 42 (1): 37 – 59.*

Kreditech (2019). Kreditech to Break Even in Cash Flow in 2020, 7 November, 2019. www.prnewswire.co.uk/news-releases/kreditech-to-break-even-in-cash-flow-in-2020-806414450.html

Kshetri, N. (2014). "The Emerging Role of Big Data in Key Development Issues: Opportunities, Challenges, and Concerns". Big Data & Society, July-December, 1-20, DOI: 10.1177/2053951714564227.

Kshetri, N. (2016a). *Big Data's Big Potential in Developing Economies: Impact on Agriculture, Health and Environmental Security*, Centre for Agriculture and Biosciences International (CABI) Publishing, Wallingford, Oxon, U.K.

Kshetri, N. (2016b)."Big Data's Role in Expanding Access to Financial Services in China," *International Journal of Information Management*, 36 (3), 297–308.
Kshetri, N., (2019). "Global Entrepreneurship: Environment and Strategy." (Routledge: New York).

Kshetri, N. (2015). "Success of Crowd-Cased Online Technology in Fundraising: An Institutional Perspective." *Journal of International Management*, 21(2), pp. 100–116.

Kshetri, N., and Acharya., S, 2012. **"**Mobile Payment in Emerging Markets." *IEEE IT Professional*, 14(4), pp. 9-14.

Lebowski, I. (2014). "How to Get a Loan without Leaving Facebook." www.wired.com/2014/05/lenddo-facebook/.

Lee, E. (2018) "WeChat claims 1.08 billion users in latest "one minute" data report," Technode, November 19, https://technode.com/2018/11/19/wechat-1-minute-statistics-report/.

Li, J.P., et. al. (2014). "Impacts of Big Data in the Chinese Financial Industry." *The Bridge* 44(4). www.nae.edu/Publications/Bridge/128772/129193.aspx; accessed 13 October 2019.

Li, X. X. (2014). "Alibaba Has Big Hopes for New Big Data Processing Service." http://english.caixin.com/2014-07-17/100705224.html; accessed 17 November 2019.

Lonsdale, R. and Archer, J. C. (1997). "Emptying Areas *in the* United States, 1990–1995." *Journal of Geography* 96, 2, 108–22.

López, M, (2020) Fintech Kueski uses AI for an end-to-end loan solution. January 27. www.contxto.com/en/mexico/fintech-kueski-uses-ai-end-to-end-loan-solutio/.

Loshin, D. (2001). "Enterprise Knowledge Management: The Data Quality Approach." *Morgan Kaufmann Publishers.*

Lunden, I. (2015). "Kreditech Nabs $92M To Build Financial Services for the 'Underbanked'." http://techcrunch.com/2015/09/28/kreditech-series-c/; accessed 11 December 2019.

Marr, B. (2016). "3 Industries That Will Be Transformed By AI, Machine Learning and Big Data in The Next Decade." *Forbes*. www.forbes.com/sites/bernardmarr/2016/09/27/3-industries-that-will-be-transformed-by-ai-machine-learning-and-big-data-in-the-next-decade/#5706fed02a11; accessed 14 December 2019.

Mason, R.O, et. al. (1997). "A Historical Method for MIS Research: Steps and Assumptions." *MIS Quarterly,* 21 (3):307-20.

McKay, C. and M. Pickens. (2010). "Branchless Banking Pricing Analysis." Consultative Group to Assist the Poor (CGAP). www.cgap.org/gm/document-1.9.43876Branchless_Banking_Pricing_Analysis_May_2010.pdf; accessed 19 September 2019.

Mckinsey Global Institute (2015). "How Advancing Women's Equality Can Add $12 Trillion to Global Growth." https://www.mckinsey.com/featured-insights/employment-and-growth/how-advancing-womens-equality-can-add-12-trillion-to-global-growth; accessed 20 March 2020.

Micco, A., Panizza, U. and Yañez, M. (2007). "Bank Ownership and Performance: Does Politics Matter?" *Journal of Banking and Finance,* 31, 219–241.

Miranda, M. J., Gonzalez-Vega, C. (2011). "Systemic risk, Index Insurance, and Optimal Management of Agricultural Loan Porfolios in Developing Countries." *American Journal of Agricultural Economics* 93(2), pp. 399-406.

Norbrook, N. and Anderson, M. (2016). "Armed with Fintech and Fueled by Data: Africa's Banker of Tomorrow." *The Africa Report*. www.theafricareport.com/Frontline/armed-with-fintech-and-fuelled-by-data-africas-banker-of-tomorrow.html; accessed 28 September 2019.

North, D. C. (1990). "Institutions, Institutional Change and Economic Performance." Cambridge, UK: *Cambridge University Press.*

North, D. C. (1996). "Epilogue: Economic Performance Through Time." In L. J. Alston, T. Eggertsson and D. C. North (Eds.). *Empirical Studies in Institutional Change*, 342–355. Cambridge University Press, Cambridge, MA.

Omidyar.com. (2015). "Asia-based FinTech startup Lenddo Raises Series B Funding," 14 October. https://omidyar.com/news/asia-based-fintech-startup-lenddo-raises-series-b-funding/.

Organization of American States. (2014). "The OAS and South-South and Triangular Cooperation: An Integral Part of our Daily Work." www.oas.org/en/media_center/press_release.asp?sCodigo=S-015/14; accessed 23 March 2020.

Parto, S. (2005). "Economic Activity and Institutions: Taking Stock." *Journal of Economic Issues*, 39 (1), 21–52.

Perez, Bien and Zen Soo (2017). "China a fast learner when it comes to artificial intelligence-powered fintech, experts say," *South China Morning Post*, October 28. www.scmp.com/tech/innovation/article/2117298/china-fast-learner-when-it-comes-artificial-intelligence-powered.

Price, R.J. and Shanks, G. (2005). "Empirical refinement of a Semiotic Information Quality framework." *System Sciences*.

Prnewswire.com. (2014a). "Accion Venture Lab Invests in Revolution Credit," May 5. www.prnewswire.com/news-releases/accion-venture-lab-invests-in-revolutioncredit-258287631.html; accessed 28 September 2019.

Prnewswire.com. (2014b). Mexican Financial Innovator Kueski.com Chooses Mambu's Banking Platform, Feb 06. www.prnewswire.com/news-releases/mexican-financial-innovator-kueskicom-chooses-mambus-banking-platform-243953481.html.

Rabinovitch, S. (2013). "Alibaba Digs Deep for Chinese Banking Treasure." www.ft.com/cms/s/0/f956b004-ffee-11e2-9c40-00144feab7de.html; accessed 26 September 2019.

Rindfleisch, Aric and Jan B. Heide. (1997). "Transaction Cost Analysis: Past, Present, and Future Applications," *Journal of Marketing,* Vol. 61, No. 4, October, pp. 30-54.

Rogaly, B. (1996). "Micro-Finance Evangelism, 'Destitute Women', and the Hard Selling of a New Anti-Poverty Formula", Development in Practice, 6 (2), pp. 100–2.

Rowley, J. (2002). "Using Case Studies in Research." *Management Research News*, 25(1), 16 – 27.

Russell, S. (1996). "Ability to Pay for Health Care: Concepts and Evidence." *Health Policy and Planning*, 11, no. 3 (1996): 219–237.

Saloom, K. (2015). "Lenddo Continues to Grow as a Social-Network Based Lending Platform for Emerging Markets, Raises an Additional US$6 Million for Global Expansion." www.prweb.com/releases/lenddo/1017/prweb11236476.htm; accessed 17 September 2019.

Schneider, S. (2013). "Five Ways Cell Phones are Changing Agriculture in Africa." http://foodtank.com/news/2013/04/five-ways-cell-phones-are-changing-agriculture-in-africa; accessed 24 November 2019.

Schwartz, P. (2014). "Peripheral Vision: Climate Change and Global Development in the 21st Century," *Pt. IV: Long-Tail Liberation*. www.huffingtonpost.com/peter-schwartz/peripheral-vision-climate_1_b_5591640.html; accessed 22 November 2019.

Schwarzmann, P. (2015). "Alibaba, Mobile, India and You: Mary Meeker and What's Trending Now." https://blog.equinix.com/2015/06/alibaba-mobile-india-and-you-mary-meeker-and-whats-trending-now/; accessed 23 January 2020.

Scidev.net. (2014). "Big Data for Development: Facts and Figures." www.scidev.net/global/data/feature/big-data-for-development-facts-and-figures.html; accessed 16 October 2019.

Seawright, J., and Gerring, J. (2008). "Case-selection Techniques in Case Study Research a menu of Qualitative and Quantitative Options." *Political Research Quarterly*, 61:2, 294-308.

Shevlin, R. (2019). "Bank/Fintech Partnerships: The Fad Is Over." *Forbes*. New York. www.forbes.com/sites/ronshevlin/2019/10/14/bank-fintech-partnerships-the-fad-is-over/#31b2eed57527; accessed 18 January 2020.

Shu, C. (2015). "Lending Platform China Rapid Finance Raises $35M Series C at A $1B Valuation," July 29. http://techcrunch.com/2015/07/29/china-rapid-finance/; accessed 17 January 2020.

Sibiko, K.W. et.al. (2018). "Small Farmers' Preferences for Weather Index Insurance: Insights from Kenya." *Agric and Food Secure*, 8, 53. https://agricultureandfoodsecurity.biomedcentral.com/articles/10.1186/s40066-018-0200-6#citeas; accessed 8 December 2019.

Smith, C. 2020 120 Alibaba Statistics and Facts (2020) | By the Numbers, July 11 https://expandedramblings.com/index.php/alibaba-statistics/

Startups.com. (2018). "Key Crowdfunding Statistics." www.startups.com/library/expert-advice/key-crowdfunding-statistics; accessed 19 January 2020.

Statistica.com. (2020). "Share of Adult population with Credit Cards in Mexico Between 2011 and 2017." www.statista.com/statistics/865785/credit-card-ownership-rate-mexico/; accessed on 13 December 2020.

Stiglitz, J.E., and A. Weiss. (1981). C"redit Rationing in Markets with Imperfect Information," American *Economic Review* 71, no. 3, pp. 393–410.

Stvilia, B., Gasser, L., Twidale, M.B. and Smith, L.C. (2007). "A framework for Information Quality Assessment." *Journal of the American Society for Information Science and Technology,* 58.

United States Senate Committee on the Judiciary, Subcommittee on Technology, Terrorism, and Government Information (2000, March 7). ID Theft: When Bad Things Happen to Your Good Name. Hearing. https://www.gpo.gov/fdsys/pkg/CHRG-106shrg69821/html/CHRG-106shrg69821.htm.

Sreedhar, S. (2015). "Financing the Supply Chain with Big Data." www.sdcexec.com/risk-compliance/article/12097108/financing-the-supply-chain-with-big-data; accessed 25 March 2020.

The, A. (2018). The retired ex-US-intel officer who started an AI business in Southeast Asia, 2 September. www.techinasia.com/exusintel-officer-retirement-start-ai-business-southeast-asia.

Un.org. (2014). "Africa Wired." www.un.org/africarenewal/magazine/special-edition-agriculture-2014/africa-wired.

UNCTAD. (2000). "Building Confidence: Electronic Commerce and Development." United Nations Conference on Trade and Development, UNCTAD/SDTE/MISC.11.

Vageesh, NS. (2016). "Tracking Online Behaviour for Granting Loans." *The Hindu BusinessLine*. www.thehindubusinessline.com/money-and banking/tracking-online-behaviour-for-granting-loans/article8991329.ece; accessed 18 August 2019.

Van Doorslaer, E. et al., (2006). "Effect of Payments for Health Care on Poverty Estimates in Eleven Countries in Asia: An Analysis of Household Survey Data." *Lancet,* 368, 1357–1364.

Van Maanen, J., Dabbs, J. M., and Faulkner, R. R. (1982). "Varieties of Qualitative Research." SAGE. Beverly Hills, CA.

Vasagar, J. (2016) "Kreditech: A Credit Check by Social Media." www.ft.com/content/12dc4cda-ae59-11e5-b955-1a1d298b6250; accessed 15 August 2019.

Voutier, P. (2019), Digital payments key to rural credit scoring in S-E Asia, 12 September, www.businesstimes.com.sg/garage/digital-payments-key-to-rural-credit-scoring-in-s-e-asia.

Wall Street and Technology (2002). "TowerGroup Finds High Costs and Lack in Consensus Main Barriers to Biometrics." http://www.wallstreetandtech.com/security/towergroup-finds-high-costs-and-lack-in-consensus-main-barriers-to-biometrics-/d/d-id/1254928?; accessed 18 December 2019.

Wang, R.Y. and Strong, D.M. (1996). "Beyond Accuracy: What Data quality Means to Data consumers." *Journal of Management information systems*, 12, 5-33.

Whetten, D. A. (1989). "What Constitutes a Theoretical Contribution? *Academy of Management Review* 14(4):490–495.

Whyte, W. F. (1984). "Learning from the Field: A Guide from Experience." SAGE. Beverly Hills, CA.

Wicem, S. (2014). "Big Data Driven Credits and Loans Startup Kreditech Raises $15M in Debt to Expand." http://technlogienews.blogspot.com/2014/01/big-data-driven-credit-and-loans.html; accessed 18 November 2019.

Wildau, G. (2015). Alibaba's Finance Arm Launches Credit Scoring Service. www.ft.com/intl/cms/s/0/34e77fe8-a6a3-11e4-9bd3-00144feab7de.html; accessed 17 August 2019.

Williamson, O. E. 1985. "The Economic Institutions of Capitalism." *The Free Press*. New York.

Williamson, O. E. (1979). "Transaction-Cost Economics: The Governance of Contractual Relations," *Journal of Law and Economics* 22(2), pp. 233-61.

Williamson, OE. (1989). "Transaction Cost Economics, Handbook of Industrial Organization," edition 1, 1(3), pp. 135-182.

Woodruff, M. (2015). "What You Need to Know About Social Media Credit Scores. http://finance.yahoo.com/news/social-media-credit-scores-145016363.html; accessed 18 January 2020.

World Bank (2018). "Financial Inclusion on the Rise but Gaps Remain, Global Findex Database Shows." Washington DC: The World Bank Group. www.worldbank.org/en/news/press-release/2018/04/19/financial-inclusion-on-the-rise-but-gaps-remain-global-findex-database-shows; accessed 15 January 2020.

World Bank. (2019). "Using Big Data to Expand Financial Services." Washington D.C.: *World Bank Group*. http://documents.worldbank.org/curated/en/505891573224492672/pdf/Using-Big-Data-to-Expand-Financial-Services-Benefits-and-Risks.pdf; accessed 23 January 2020.

Xinhua, (2016). "China Tops Global Fintech Rankings: Report." *China Daily*, 7 November. http://usa.chinadaily.com.cn/business/2016-11/07/content_27298823.htm; accessed 13 January 2020.

Yin R. (1994). "Case Study Research: Design and Methods" (2nd ed.). SAGE. Newbury Park, CA.

Zendesk.com. (2017). "Kreditch Serves the Underbanked with Zendesk." www.zendesk.com/customer/kreditech/; accessed 23 November 2019.

Zerucha, T. (2019). "Kueski building financial inclusion in Latin America." 8 July. www.banklesstimes.com/2019/07/08/kueski-building-financial-inclusion-in-latin-america/.
Zetsche, D., Buckley, R., Arner, D., and Barberis, J. (2018). "From Fintech to Techfin: The Regulatory Challenges of Data-Driven Finance." *New York University Journal of Law and Business*, 14, 393-446.

Zhang, S. Woo, R (2018). "Alibaba-backed online lender MYbank owes cost-savings to home-made tech," 31 January. www.reuters.com/article/us-china-banking-mybank/alibaba-backed-online-lender-mybank-owes-cost-savings-to-home-made-tech-idUSKBN1FL3S6.