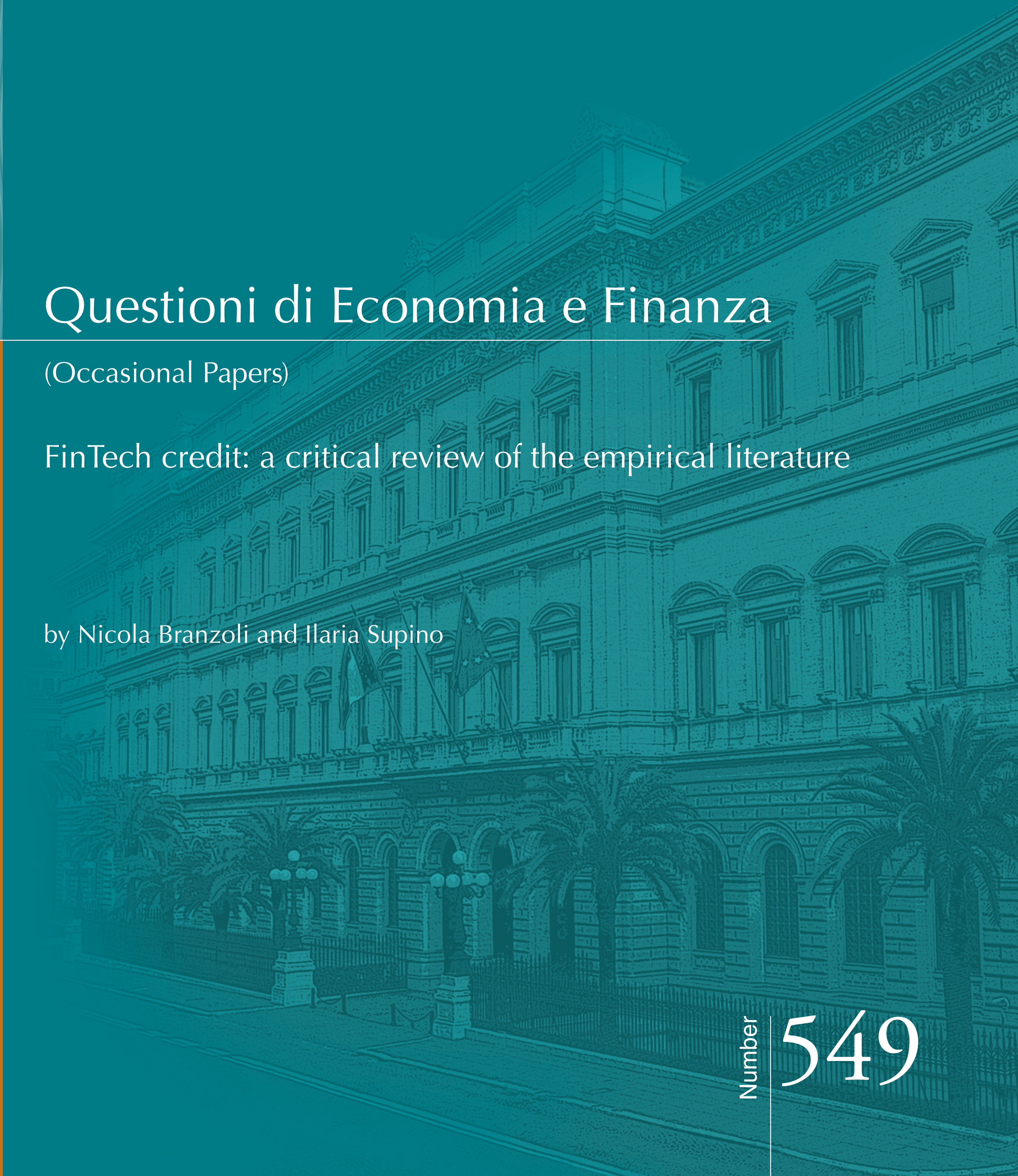
BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

FinTech credit: a critical review of the empirical literature

by Nicola Branzoli and Ilaria Supino



March 2020

Number 549

BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

FinTech credit: a critical review of the empirical literature

by Nicola Branzoli and Ilaria Supino

Number 549 – March 2020

The series Occasional Papers *presents studies and documents on issues pertaining to the institutional tasks of the Bank of Italy and the Eurosystem. The* Occasional Papers *appear alongside the* Working Papers *series which are specifically aimed at providing original contributions to economic research.*

The Occasional Papers *include studies conducted within the Bank of Italy, sometimes in cooperation with the Eurosystem or other institutions. The views expressed in the studies are those of the authors and do not involve the responsibility of the institutions to which they belong.*

The series is available online at www.bancaditalia.it.

ISSN 1972-6627 (print)
ISSN 1972-6643 (online)

Printed by the Printing and Publishing Division of the Bank of Italy

FINTECH CREDIT: A CRITICAL REVIEW OF THE EMPIRICAL LITERATURE

by Nicola Branzoli[*] and Ilaria Supino[*]

Abstract

FinTech credit has attracted significant attention from academics and policymakers in recent years. Given its growing importance, in this paper we provide an overview of the empirical research on FinTech credit to households and non-financial corporations (NFCs). We focus on three broad topics: i) the factors supporting the development of innovative business models for credit intermediation, such as marketplace lending; ii) the benefits of new credit risk assessment data and methods; iii) the implications of these innovations for access to credit. Three main messages emerge from the literature. First, the growth of lenders with innovative business models is mainly driven by the degree of local economic development and of competition in the banking sector. Second, new data and methods can improve traditional credit risk models because they are particularly helpful in screening opaque borrowers, such as those with scant credit history. Third, FinTech borrowers generally lack (or have limited) access to finance and tend to be riskier than traditional bank borrowers.

JEL Classification: G21, G22, G23, G24.
Keywords: artificial intelligence, credit, digital technologies, FinTech, marketplace lending.
DOI: 10.32057/0.QEF.2020.549

Contents

[*] Bank of Italy, Financial Stability Directorate, Directorate General for Economics, Statistics and Research.

1 Introduction *

Financial technology, or FinTech for short,[1] has progressed rapidly over the past few years. The surge in the use of internet and other digital advancements have helped introduce a variety of innovations in the financial sector forcing market incumbents to rethink their business models, with important consequences for the future of the industry as a whole (McKinsey (2019); Petralia *et al.* (2019)).

While the most noticeable change occurred in the segment of payments, FinTech companies have then gradually moved into core banking services, including the credit business. Like banks, FinTech operators provide consumer lending, corporate financing, and grant mortgage loans. Credit provided by or via FinTech companies continues to expand at a fast pace, though remaining limited relative to credit extended by traditional intermediaries (Claessens *et al.* (2018)).

To date, research efforts directed to organize the growing literature on topics that fall under the label of FinTech have been either very broad (see Thakor (2019) for a general overview) or, vice versa, focused on a specific aspect of the phenomenon (for example, Morse (2015) on peer-to-peer lending). Our work complements these recent efforts and offers a novel, critical assessment of the empirical literature on the development and functioning of FinTech credit markets.

The review presented here is organized around three topics. First, we investigate the drivers underpinning the rise of FinTech credit; based on the evidence gathered in the literature review, we identify and discuss both demand-side and supply-side factors that explain the expansion of FinTech credit. Second, we synthesize recent research on the use of new data and techniques made available by technological improvements in the financial industry, devoting specific attention to understand whether and to what extent Big Data tools allow lenders to overcome their informational disadvantages to the borrower and to perform better credit risk assessment. Third, we focus on the studies that explore the role of FinTech lending as an enabler for broader and better access to financing.

*We thank Francesco Columba, Giuseppe Ferrero, Paolo Finaldi Russo, Giorgio Gobbi, Giovanni Guazzarotti and Silvia Magri for their comments. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Italy.

[1] There is no agreed definition of the term FinTech. For the purposeof this work, we use the one suggested by the Financial Stability Board (Financial Stability Board (2017)), which defines FinTech as "technology enabled innovation in financial services that could result in new business models, applications, processes or products with an associated material effect on the provision of financial services".

Based on these three areas of research, we highlight the following insights: (i) FinTech credit grows more in wealthier regions and where bank competition is lower; (ii) the availability of alternative data (other than standard information) improves access to finance for opaque borrowers; (iii) risk assessment models based on new techniques that perform better than traditional ones; (iv) FinTech borrowers are riskier than bank borrowers.

All in all, results derived from the review suggest that the inception of FinTech players into credit markets could yield a number of benefits, mostly in terms of enhanced financial inclusion and reduced information asymmetries. However, we caution that the actual effects related to the use of new technologies for credit provision are still unclear and have yet to be tested through a full financial cycle.

The rest of the paper is structured as follows. Section 2 analyses the evidence on the drivers of FinTech credit across the world. Section 3 summarizes the literature on the implications of using new data and methods for credit risk assessment. Section 4 reviews research on the effects of technological innovation on access to credit and borrowers' selection. Section 5 concludes.

2 Drivers of FinTech credit

A growing literature focuses on the drivers of FinTech credit analyzing the development of marketplace lending, i.e. credit activity facilitated by electronic platforms. Platforms' business model is based on the disintermediation of credit activity through the direct matching of lenders with borrowers.[2] There are two main categories of marketplace lending, one in which lenders are retail consumers (peer-to-peer lending, P2P) and one in which lenders are wholesale investors (banks or institutional investors).[3] Most platforms combine both types of lenders and specialize either in consumer credit or in credit to NFCs, mostly small and medium enterprises.

Three main insights emerge from this literature. First, like traditional financial intermediation, FinTech credit is mainly driven by local economic development and the presence of a legal system that protects property rights and ensures the rule of law. A second, and somewhat surprising, set of results suggests that technological variables, such as the diffusion of Internet among the population or the presence of

[2] Marketplace lending is also based on the reduction of borrowers search costs and on the use of Internet rather than physical branches for the supply of loans, which could greatly reduce the costs of credit intermediation. However these characteristics are not related only to marketplace lending. The reduction of search costs is obtained also through the diffusion of price comparison websites, i.e. websites that compare loan offers by different banks; Internet-based credit intermediation is also related to the development of so-called neobanks, i.e. banks that operate only through Internet.

[3] See Bofondi (2017) for a description of all potential business models of marketplace lending.

broadband connections, are not significant drivers of FinTech credit growth. Third, FinTech credit is more developed where the banking sector is less competitive.

These results have broad implications for policymakers. They seem to suggest that fostering local economic development and guaranteeing a well-functioning legal system can be relatively more important than investing in high quality technological infrastructures to promote innovation in credit markets. Furthermore, FinTech credit can provide significant benefits in terms of competition in the traditional banking sector.

In the next two paragraphs we discuss in details this literature, organizing the presentation around demand and supply factors.

2.1 Demand factors

Cross-country studies that measure the development of marketplace lending using the amount of credit originated through platforms[4] find that local economic activity (e.g. income, the level of unemployment) explain a large fraction of the heterogeneity in marketplace lending (Claessens *et al.* (2018), Rau (2018)). Jagtiani and Lemieux (2018a) and Tang (2019) find similar evidence using data in the consumer credit market from LendingClub, one of the main P2P platform in the US; Buchak *et al.* (2018) and Haddad and Hornuf (2018) confirm this result using, respectively, loan level data from the U.S. mortgage lending market and investments in FinTech startups specialized in credit.

The diffusion of Internet among consumers is another potential factor underpinning Fintech expansion commonly investigated in the literature. Internet use is found to be a key driver of the demand of FinTech services in many surveys (EY (2016)); however, surprisingly, the analyses surveyed in this review do not support this expected significance for the industry.

Rau (2018) provides evidence that the percentage of the population using Internet does not help explaining cross-country differences in the volume of marketplace lending per capita. Using data from the US mortgage market, Fuster *et al.* (2019) find that lack of access to Internet or the diffusion of broadband among the population are not significant drivers of the development of FinTech mortgage lending across US areas (more precisely, census tracts).

[4]Generally normalized by country's population or gross domestic product.

Finally, it is sometimes argued that the lack of consumer trust in the banking system might favor the development of alternative sources of financing such as FinTech credit (CGFS and FSB (2017)). None of the analyses reviewed in this survey finds that proxies of consumer trust in the banking system help to explain the variability of marketplace lending across countries.

2.2 Supply factors

Supply factors underpinning the expansion of platform-intermediated credit are often categorized into three broad groups (CGFS and FSB (2017)). The first group is related to the development of the legal system. The strength of borrowers' and lenders' legal protection and the ability of public authorities to enforce their rights influence the development of traditional credit markets (Porta *et al.* (1998)); therefore many studies hypothesize that they might also play a key role in the growth of FinTech credit. The development of the legal system, proxied by various measures of the quality of contract enforcement, of the protection of property rights and of lender rights by bankruptcy laws, are generally found to be highly significant (Rau (2018), Haddad and Hornuf (2018)).[5]

The second group of factors is related to the level of technological advancement in a country or a region. For example, the presence of fast Internet connections can improve the quality of services offered by FinTech firms and increase their ability to poach consumers from incumbent banks.

Technological factors, though, do not seem to play a significant role in the development of FinTech credit. Haddad and Hornuf (2018) include an indicator of technological advancement developed by the World Bank and do not find it significant in explaining the heterogeneity in the number of FinTech start-ups specialized in credit. Fuster *et al.* (2019) present a case study of the roll-out of Google Fiber in Kansas City in 2012 and find no impact on the development of technology-based mortgage lending.

Finally, a third group of factors is related to the degree of competition and to the stringency of regulation in the banking sector. While more competition is found to deter the development of FinTech credit because it reduces expected profits for start-ups (Claessens *et al.* (2018)), the effect of regulation within the banking sector perimeter is not clear *a priori*. On the one hand, regulation can hinder FinTech credit

[5] Analyses using data from single countries do not address this issue as their samples do not contain variation in the strengths of legal rights.

because some regulatory costs are, from an economic perspective, a type of entry costs, reducing the incentives for firms that want to offer FinTech credit. On the other hand, higher regulatory costs may reduce the supply of loans by traditional financial institutions, making it easier for FinTech firms to gain market shares.

Most studies include proxies for the degree of competition in the banking sector, finding that marketplace lending is more developed where the banking sector is less competitive. Claessens *et al.* (2018) and Rau (2018) find that the Lerner index in the banking sector and the concentration of banks' assets in the top five intermediaries, which are a common indicators for the lack of competition, are positively associated with per capita platform-intermediated credit in a country. Frost *et al.* (2019) confirm this evidence focusing on the amount of credit originated by large technology companies, c.d. BigTechs.[6] Using data on consumer loans that are used to pay off credit card balances or for debt consolidation from the P2P platform Lending Club, Jagtiani and Lemieux (2018a) examine in which areas the lending platform expanded the most. Their results indicate that new loan origination from Lending Club is higher in areas where the market for credit card loans is more concentrated. Buchak *et al.* (2018) show that platform-intermediated credit is more developed in counties where there is less banking competition. Rau (2018) finds that the cost-income ratio in the banking sector, a proxy of its inefficiency, is positively related to the development of marketplace credit.

There is no univocal evidence on the relation between the stringency of banking regulation and the development of marketplace lending. Claessens *et al.* (2018) find that the regulatory stringency index for the banking sector constructed by Navaretti *et al.* (2017) is negatively associated with the development of FinTech credit. However, Braggion *et al.* (2018) finds that tightening banks' macroprudential regulation cause an increase in the volumes of credit intermediated by platforms.

[6]The term "BigTech" is generally used to describe the direct provision of financial services or products by technology companies, such as Apple, Google, Amazon, Facebook, Ant Financial, Ali Baba, Mercado Libre, Vodafone and Samsung (Claessens *et al.* (2018)).

3 New data and techniques for risk assessment

Digital innovation has changed the amount and the nature of the information used for lending decisions. With the development of e-commerce and social media, the amount of data that can be used for credit assessment has increased considerably and technological advancements have lowered the cost of storing and processing this information.

In the traditional approach to risk assessment, banks mainly use standardized data sources, such as credit registers and credit bureaus, which in some cases are integrated with soft information about the borrower. Big Data[7] and Artificial Intelligence[8] strengthen the benefits of proactive information gathering, an activity which FinTech firms tend to do very efficiently given their ability in handling digital inputs. Lenders can complement standard sources with new data that can be scraped from rating websites, price comparison websites, networking platforms and online marketplaces in order to build more comprehensive credit profiles of prospective borrowers, even on a daily basis, increasing the frequency of their monitoring.

In this section we review studies on the impact of digital technologies on different outcomes of the lending decision-making process. Two main insights emerge: (i) online data additional/alternative to traditional ones can help screen borrowers, especially those with poor credit history; (ii) model based on innovative techniques perform better than traditional models in predicting loan performance.

3.1 Evidence on the use of alternative data sources

The use of online data to assist lending decisions and, more specifically, credit risk management has attracted considerable academic interest. Berg *et al.* (2018) analyze the use of digital footprints (i.e. the information left behind by individuals while navigating the Internet) to gauge the creditworthiness of potential customers. Using a comprehensive dataset covering roughly 270,000 purchases at a German

[7] The term 'Big Data' generally refers to datasets that are too large to be analyzed with traditional statistical techniques. Einav and Levin (2010) identify three key features that distinguish Big Data from traditional datasets. First, Big Data are usually updated in real-time. Second, they describe activities that were previously difficult to quantify, such as personal communications, social networks, search and information gathering, geolocation data. Finally, Big Data do not have a clear and unique structure, but can be organized in multiple ways. For example, Internet browsing histories contain a large amount of information about a person's interests and beliefs and how they evolve over time.

[8] Artificial Intelligence is a branch of computer science that studies how to develop devices that perceive the environment where they operate and take actions to maximize the chance of successfully achieving certain goal(s).

e-commerce company between October 2015 and December 2016, the authors compare the predictive power of digital footprint variables (proxying for income, character, and reputation) vis-á-vis traditional credit bureau scores.[9] The paper finds that risk assessment models that use only the digital footprint variables equal, and in some cases exceed, the information content of credit bureau scoring. One of the main implications of this results is that when credit bureau information is not available, as for example for borrowers with little or no credit history, digital footprints might help lenders to mitigate problems of information asymmetries, thus alleviating credit constraints for borrowers without credit scores. Models that combine information collectable from digital and traditional sources display discriminatory power greater than traditional models. The literature finds that the use of information extracted from payment data (for example using credit card information) in credit scoring models improves their performance (Wilson *et al.* (2000), Khandani *et al.* (2010)). Frost *et al.* (2019) compare the predictive content of the internal rating developed by Mercado Libre, an Argentine company that operates online marketplaces dedicated to e-commerce and online auctions, for small and medium enterprises[10] with that of local credit bureau rating. The authors find that the internal rating developed by the Argentinian BigTech allows borrowers rated as medium and high risk by the local credit bureau rating to be correctly assessed as low risk (so called "undetected primes"). Gambacorta *et al.* (2019) show that models combining traditional bank data with non-traditional information obtained from credit card data, digital applications and e-commerce platforms perform better than models that use only the former. In a similar vein, Iyer *et al.* (2016) examine how lenders in P2P markets use non-standard information in screening borrowers. Their sample contains approximately 200,000 listings posted on a large US P2P platform (Prosper.com) between February 2007 and October 2008. For each loan listing, the authors observe both the traditional credit score[11] and non-standard variables such as the maximum rate that the applicant is willing to pay or personal text descriptions (soft information available to online lenders). The analysis aims at understanding whether they outperform the credit score system in terms of default

[9]The quality of each default prediction model is measured as the Area Under a Receiver Operating Characteristics (AuROC) curve, a common metric used to assess the accuracy of a screening test in a setting with a continuous predictor for a dichotomous outcome. AuROC curves are estimated using logit regressions with a default dummy as the dependent variable.

[10]Mercado Libre's internal rating model uses more than a thousand variables based on sale volumes, average selling prices and seller's reputation (e.g. number of complaints).

[11]Credit scores are unknown by the peer lenders, who only see the aggregate credit category.

predictive power.[12] The interest rate required to partially or fully fund a loan application - which the authors interpret as lenders' inference on borrowers' risk - is decomposed along different information sources.[13] Interest rates set by online lenders are found to predict default more accurately than traditional credit score, with non-standard information being relatively more important when assessing bad quality borrowers.

Jagtiani and Lemieux (2018b) look into loan-level data from LendingClub and large US banks in order to compare loans originated by traditional intermediaries vis-á-vis those extended by FinTech lenders. They find that credit provided by FinTech operators is cheaper than comparable credit granted via traditional lending channels, and that it is priced accurately given expected default rates. The adoption of non-standard information to evaluate loan applications is shown to have benefited pools of under-served borrowers (those with fewer or partial credit records).

A parallel stream of research has started to analyze the informative content of narratives attached to credit requests. Online loan applications are typically accompanied by text paragraphs where borrowers provide additional personal information in order to increase their chances of being funded. Herzenstein *et al.* (2011) examine the role played by narratives in facilitating credit extension in online marketplaces. They find that, on the one hand, claiming trustworthy identities results in larger and cheaper loan funding; on the other hand, good borrower's self representations are not predictive of a positive loan performance.

A related point encompasses the importance of social interconnections in predicting online lending outcomes. Lin *et al.* (2013) test for the role of friendship in driving successful funding applications. Examining ex-post outcomes of transactions on a P2P marketplace, the study finds that borrowers with friends logged over the same platform experience higher loan acceptance rates and face lower prices. The information content of a friendship tie acts as a signal for a borrower's credit quality and helps online lenders to screen better counterparties. However, Freedman and Jin (2018) suggest caution for

[12]The authors also assess whether non-expert market participants aggregate and process listing information to infer borrowers' creditworthiness.

[13]Methodologically, the quality of the screening process is measured as both the goodness-of-fit (R2) from a linear regression of ex-post loan performance on predictors (either the credit score or the interest rate) and - similar to Berg *et al.* (2018) - as the area under the receiver operating characteristics (ROC) curve. Authors construct four ROC curves: one for the interest rate, one for the credit score, one for all the standard financial variables, one for all available (standard and nonstandard) variables. The larger the area under the curve the better the predictive power of the screening test.

using online social networks to predict positive loan performance, showing that borrowers who obtain easier and cheapest credit access thanks to their social ties are more likely to delay payments or default.

3.2 Evidence on the use of innovative methods for credit scoring

The combination of standard and alternative information sources often creates large amounts of data which are hard to analyze with standard credit scoring models like linear and logit regressions. A growing literature has started to address this issue by investigating the benefits of applying Artificial Intelligence and Machine Learning (AI-ML) techniques for credit risk assessment. These innovative methods can handle dataset with thousands of variables for each borrower in a non-parametric way, taking into account non-linear complex functions of these variables. This flexibility have both benefits and pitfalls. These models tend to provide better out-of-sample forecasts than traditional models but the relationship between the dependent and the explanatory variables is often hard to interpret from an economic point of view (s.c. "black box"[14] problem).[15]

AI-ML models tend to outperform traditional credit models when few borrower characteristics are available, thanks to their better ability to extract information from non-linear relationships between the variables. Moscatelli *et al.* (2019) show that ML models have better out-of-sample performance when the models are estimated[16] using only public information; they display roughly the same performance of traditional models when data from the credit register are available. Gambacorta *et al.* (2019) show that the benefits of using credit scoring models based on ML instead of traditional models increase when the relationship between the lender[17] and the borrower is shorter.[18] This evidence suggests non-traditional models can improve the risk assessment for borrowers with little available information, such as those with a short credit history.

As with the use of alternative data, one of the main benefits of using new methods is related to their superior ability to find undetected primes.[19] Albanesi and Vamossy (2019) develop a credit scoring

[14]Bracke *et al.* (2019) provides a discussion of this issue and proposes a framework to solve it.
[15]See Atiya (2001) for a early survey of analyses predicting bankruptcy with ML models.
[16]In machine learning, the term "trained" is often used instead of "estimated".
[17]In their application, the lender is a leading FinTech company in China.
[18]Lenders have less information available about the borrowers when their relationship is shorter.
[19]See the previous section for a brief characterization of these borrowers.

model based on deep learning that significantly outperforms the traditional logistic model.[20] The authors show that this improvement benefits mainly consumers with low credit scores: the deep learning model is able to identify among them the borrowers that are less likely to default. For such borrowers they also quantify the credit card interest rate savings of being classified according to the deep learning model in about five percent of the debt balance (or one quarter of the average interest rate expense of low credit score borrowers).

A key question is related to the performance of AI-ML models during economic downturns. In such conditions the relationship between the variables may change and AI-ML models, which to be estimated require significant amount of data, may be slow to capture such changes.[21] Gambacorta *et al.* (2019) analyze the performance of ML models before and after a regulatory shock in China that restricted funding for shadow banks and resulted in a negative credit shock to the Chinese economy and in an increase in the number of defaults. The paper finds that the performance of both traditional and non-traditional models deteriorate after the shock, but less for the latter. The authors attribute this result to the non-linearity of the ML models, which may better capture dynamic relationships that are more relevant when the external environment suddenly changes.

Credit scoring models are often compared at different forecast horizons. Given similar performance at short horizons (i.e. 30 or 60 days ahead), models that provide better forecasts at longer horizons may allow lenders to intervene earlier if borrower's quality deteriorate and to prevent losses. Butaru *et al.* (2016) show that when credit scoring models are used to predict defaults one or two quarters ahead, ML-based models still outperform traditional models, but the relative benefits are smaller than those at shorter horizons. This is likely due to the tendency of ML-based models to overfit the sample with which they are estimated.

One topic often discussed in the literature is whether the use of new methods help reduce minority

[20]The model is estimated using standard data from consumer credit reports (such as debt balances and number of card operations) and it is used to forecast delinquencies of consumer loans within a 3 months horizon. They also show that the default probability predicted by their approach is slightly better than the credit score in predicting actual defaults. The performance of the deep learning model and the credit score is compared using two measures. The first is the rank correlation between realized default rates and the credit score or between realized default rates and the probability of default based on the deep learning. The second is the Gini coefficient, which measures the dispersion of the credit score distribution and therefore its ability to separate borrowers by their default risk.

[21]This issue is often referred to as "overfitting". A model is said to overfit the data if it describes closely or exactly the dataset on which it is trained but fails to fit other data or predict future observations reliably (out-of-sample prediction).

discrimination in the credit market. Dobbie *et al.* (2018) document the presence of considerable racial bias in the UK consumer lending market, showing that a decision rule based on ML predictions could eliminate such bias. Fuster *et al.* (2018) study the potential implications for different race and ethnic groups of using machine learning models for credit risk assessment in the US mortgages market. They find that changes in predicted default propensities across race and ethnic groups differ significantly. A large fraction of borrowers belonging to the majority group (e.g., white non-Hispanic) experience lower estimated default propensities under the machine learning ML technology than under the logit model. This group can potentially benefit from lower average interests rates. Other groups (e.g., black and Hispanic borrowers) do not accrue to the same level of improvement, therefore the change in the average interest rate is smaller. Borrowers in these groups mainly experience an increase in the dispersion of rates, as some borrowers benefit from the greater default accuracy and some do not.

Finally, given that this literature is relatively new, there is significant uncertainty about which subclass of AI-ML models provides better performance. Baesens *et al.* (2015) review the recent literature on credit scoring and recommends a particular AI technique called random forests.[22] Albanesi and Vamossy (2019) show that models based on deep learning can provide better performances than random forests in predicting credit default. Focusing on credit card delinquencies, Butaru *et al.* (2016) find that decision tree models tend to perform better than those based on deep learning.

4 FinTech credit: access and borrowers' characteristics

FinTech lending is often seen as an innovative avenue for expanding access to credit and facilitating market participation in segments with no (or scarce) financing opportunities. A growing number of studies examine credit allocation by FinTech lenders to assess whether shifts in the composition of borrowers (towards those who were either unserved or underserved by traditional banks) occur and whether such changes come along with a deterioration in the quality of loan portfolios.

In this section we describe two main sets of findings. The first group of results concern the issue of financial inclusion, with a common assumption being that FinTech lending helps fill the credit gap

[22] Random forests are a set statistical techniques based on Artificial Intelligence used to cluster observations into homogeneous groups.

for borrowers that lack (or have limited) access to finance. The second set of findings relates to the characterization of FinTech borrowers in terms of riskiness.

We draw the following conclusions: (i) FinTech expands credit access for marginal borrowers with credit demand unmet by traditional borrowers; (ii) FinTech borrowers are riskier than traditional ones.

4.1 Financial inclusion

The idea that FinTech lenders serve unmet customer demand is grounded in an emerging body of research that focuses specifically on credit markets in developing countries. Hau *et al.* (2019) show that online lending extends the frontier of credit availability to firms with low credit scores. Firms use FinTech credit to fund growth (Hau *et al.* (2018)) and expand their product offering (Frost *et al.* (2019)). Hau *et al.* (2019) model the entry of a FinTech firm into the Chinese credit segment for small businesses. Authors conjecture that FinTech lending increases the extensive margin of credit to vendors with larger ex-ante credit risk, who were previously excluded by the traditional banking system. The theoretical prediction is tested empirically using account-level data from the largest provider of automated online credit between 2014 and 2016: FinTech credit demand is found to be larger in urban areas with less bank credit supply (relative to local GDP) and in rural areas with greater distance between the loan applicant and the nearest bank branch.[23]

In developed countries, the literature generally finds marketplace credit is often used for refinancing existing bank loans, indicating that banks and platforms mainly target the same kind of borrowers. Buchak *et al.* (2018) show that in the mortgage lending market a refinancing loan is 20 percent more likely to be granted from an online platform rather than from a bank or a shadow bank. Bayluk (2018) documents that more than 80 percent of the volumes of loans originated through Prosper are used to consolidate bank credit card debts. The competition between bank and marketplace credit in refinancing is confirmed by the evidence that the latter tends to develop more in areas where banks are hit by a negative cost shocks and reduce credit (Tang (2019), de Roure *et al.* (2018)).

The development of marketplace lenders partly crowds out loans issued by banks (de Roure *et al.*

[23] Similar results have been found for US by Jagtiani and Lemieux (2018a), who find that FinTech consumer lending have penetrated areas where the number of bank branches has decreased more than others.

(2018)). However this crowding-out effect is not the same across all banks. Wolfe and Yoo (2018) combine data from Prosper and Lending Club to show that small commercial banks (with assets below 300 million) have reduced lending volumes in the consumer lending market in response to the entry of P2P platforms, while other commercial banks do not appear to have been impacted by the competition of P2P lending.[24]

FinTech's role in boosting financial inclusion is less clear. For the US mortgage market, Bartlett *et al.* (2019) find that FinTech lenders are much less discriminatory (than their face-to-face peers) in loan approval decisions. For the US consumer market, however, Duarte *et al.* (2012) document disparate treatment based on perceptual factors that could affect lenders' decisions within the online marketplace. Using photographs that customers uploaded on a P2P lending website, authors test whether appearance-based impressions inform investor behavior in terms of funding choices. For each potential borrower, a pool of independent raters was asked to rate the willingness to pay associated with the person in the picture posted online. Individual judgments were then averaged to construct a measure of the borrower's perceived trustworthiness. Their results show that borrowers who appear more trustworthy have greater probabilities to secure a loan (at relatively lower interest rates), exhibit better credit scores and default less frequently.

Recent research suggests that the likelihood of funding success on FinTech credit markets also varies across racial and social groups. Pope and Sydnor (2011) find that loan listings featuring photographs of black people are significantly less likely to receive full funding. They also demonstrate that racial discrimination is reflected in the interest rates black-skin borrowers pay conditional on obtaining a loan, these rates being sensibly higher than those charged to white peers with identical credit profiles. Duarte *et al.* confirm that lender attitude toward ethnicity is an important predictor of funding. Their results suggest taste-based discrimination against minority borrowers whose loan request listings are, on average, less likely to get approved. Online lenders are also found to discriminate borrowers on the basis of demographic attributes such as age (Pope and Sydnor (2011); Gonzalez and Komorova Loureiro (2014)) or attractiveness (Ravina (2012)), while the effect of gender differences is still not

[24] To identify the influence of P2P lending, the authors utilize time varying, state level entry restrictions on the part of P2P borrowers and P2P investors and an instrumental variable strategy based on the fraction of the population able to supply funds on the P2P lending platforms.

clear (Potzsch and Bohme (2010); Pope and Sydnor (2011)).

4.2 The riskiness of FinTech borrowers

A number of empirical studies show that online lenders tend to target more financially vulnerable individuals and firms. In the consumer credit segment, Maggio and Yao (2019) study the features of the pool of households that borrow from FinTech credit providers in comparison with their peers served by traditional banks. Using panel data containing detailed information on individual credit profiles, authors find that - at loan origination - FinTech borrowers are on average high-income young professionals with good credit histories who become new customers for FinTech lenders (but not because they are shut out of the traditional banking system) in order to obtain additional, cheaper credit. This ex-ante high creditworthiness, however, does not result in a better ex-post performance since Fintech loans are found to be (3 percent) more likely to default in the months following the origination: the underlying mechanism driving this result - argue Maggio and Yao (2019) - is the need for immediate consumption which leads households to overborrow and poorly perform on their loans.

The higher riskiness of FinTech borrowers is also documented in Tang (2019). Exploiting a regulatory change as an exogenous shock to bank credit supply, the author analyses changes in the distribution of FinTech borrowers quality under the assumption that tighter credit standards induce riskier, rationed bank clients to migrate to peer-to-peer platforms. The paper shows that, in the markets exposed to bank credit supply shock, the quality of the Fintech borrower pool deteriorates as bank credit availability decreased (at the expense of more vulnerable individuals), suggesting that FinTech lending substitutes traditional lending only for infra-marginal bank clients. Similarly de Roure *et al.* (2018) find that FinTech lenders do not 'cream-skim' better borrowers, but rather they cater for the riskiest segment of bank clients.

In the residential mortgage market, many papers that detect borrowers' differences across lender types confirm that FinTech credit flows toward riskier costumers. Buchak *et al.* (2018) show that shadow banks (which include FinTech players) specialize in lending to less creditworthy borrowers (relative to traditional banks). Allen *et al.* (2019) compare risk-taking across bank and FinTech lenders in the event of a natural disaster: while loan risk attributes improve in the wake of local demand shocks for traditional banks (suggesting a relative tightening of lending standards), the riskiness of mortgages originated after the disaster by FinTech lenders remain unchanged (with no shift toward better quality borrowers).

While most of the papers presented above support the view of FinTech operators as higher risk-taking, there are a few studies that point in the opposite direction. For example, Fuster *et al.* (2019) use data on mortgage applications and originations in US to document the superior performance of Fintech loans and reject the hypothesis that FinTech lenders 'lax-screen' borrowers, selecting risky or marginal ones. This is in line with Jagtiani and Lemieux (2018a) who find no evidence that consumers borrowing from a peer-to-peer platform in US represent the lower end of creditworthy consumers.

5 Conclusions

Banks fulfill two critical functions in credit markets: they create loans from deposits and they reduce asymmetric information frictions by selecting and monitoring borrowers. Digital technologies are radically changing both functions. Internet platforms can directly match savers and borrowers, Big Data and Artificial Intelligence are powerful technologies to mitigate asymmetric information issues. We reviewed the empirical literature on the effects of these technologies in credit markets. We found evidence that supports the idea that they can bring competition in credit markets and foster financial inclusion, particularly for opaque borrowers with little credit history. We conclude by pointing to an issue that, in our opinion, has been analyzed less than it deserves. There is little and conflicting evidence on whether these technologies help lenders to align interest rates with borrowers' risk, i.e. if they can improve the pricing of risk in the economy. For example, platform-intermediated credit is generally cheaper but sometime riskier than bank credit, suggesting that marketplace lenders may underprice borrowers risks. Big Data and Artificial Intelligence can reduce asymmetric information between borrowers and lenders, but they can also improve the ability of lenders to price discriminate borrowers and charge higher interest rates to those with higher willingness to pay. These contrasting effects highlight the need for more research in this area.

References

ALBANESI, S. and VAMOSSY, D. F. (2019). Predicting consumer default: A deep learning approach. *NBER Working Paper*, (26165).

ALLEN, L., SHAN, Y. and SHEN, Y. (2019). Do fintech lenders fill the credit gap? evidence from local mortgage demand shocks. *Working Paper*.

ATIYA, A. F. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE TRANSACTIONS ON NEURAL NETWORKS*, **12** (3), 929–935.

BAESENS, B., LESSMANN, S., SEOW, H.-V. and THOMAS, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, **1** (247), 124–136.

BARTLETT, R. P., MORSE, A., STANTON, R. H. and WALLACE, N. E. (2019). Consumer-lending discrimination in the fintech era. *NBER Working Paper*, (25943).

BAYLUK, T. (2018). Financial innovation and borrowers: Evidence from peer to peer lending. *Rotman School of Management Working Paper*, (2802220).

BERG, T., BURG, V., GOMBOVIĆ, A. and PURI, M. (2018). On the rise of fintechs - Credit scoring using digital footprints. *NBER Working Paper*, (24551).

BOFONDI, M. (2017). Lending-based crowdfunding: opportunities and risks. *Bank of Italy Occasional Papers*, (375).

BRACKE, P., DATTA, A., JUNG, C. and SEN, S. (2019). Machine learning explainability in finance: an application to default risk analysis. *Bank of England Staff working paper*, (816).

BRAGGION, F., MANCONI, A. and ZHU, H. (2018). Can technology undermine macroprudential regulation? Evidence from peer-to-peer credit in China. *SSRN Working Paper*, (2957411).

BUCHAK, G., MATVOS, G., PISKORSKI, T. and SERU, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics*, (130), 453–483.

BUTARU, F., CHEN, Q., CLARK, B., DAS, S., LO, A. W. and SIDDIQUE, A. (2016). Risk and risk management in the credit card industry. *Journal of Banking & Finance*, **72**, 218–239.

CGFS and FSB (2017). Fintech credit: Market structure, business models and financial stability implications. *May*.

CLAESSENS, S., FROST, J., TURNER, G. and ZHU, F. (2018). Fintech credit markets around the world: size, drivers and policy issues. *BIS Quarterly Review*, (September).

DE ROURE, C., PELIZZON, L. and THAKOR, A. V. (2018). P2P lenders versus banks: Cream skimming or bottom fishing? *SSRN Working Paper*, (3174632).

DOBBIE, W., LIBERMAN, A., PARAVISINI, D. and PATHANIA, V. (2018). Measuring bias in consumer lending. *NBER Working paper*, (24953).

DUARTE, J., SIEGEL, S. and YOUNG, L. (). To lend or not to lend: Revealed attitudes towards gender, ethnicity, weight, and age in the u.s. *SSRN Working paper*.

—, — and — (2012). Trust and credit: The role of appearance in peer-to-peer lending. *The Review of Financial Studies*, **5** (8), 2455–2483.

EINAV, L. and LEVIN, J. (2010). Economics in the age of big data. *Science*, **346** (6210).

EY (2016). The EY fintech adoption index: Exploring a new financial services landscape. *January*.

FREEDMAN, S. and JIN, G. Z. (2018). The information value of online social networks: Lessons from peer-to-peer lending. *International Journal of Industrial Organization*, **51**, 185–222.

FROST, J., GAMBACORTA, L., HUANG, Y., SHIN, H. S. and ZBINDEN, P. (2019). Bigtech and the changing structure of financial intermediation. *BIS working paper*, (779).

FUSTER, A., GOLDSMITH-PINKHAM, P., RAMADORAI, T. and WALTHER, A. (2018). Predictably unequal? the effects of machine learning on credit markets. *SSRN Working Paper*, (3072038).

—, PLOSSER, M., SCHNABL, P. and VICKERY, J. (2019). The role of technology in mortgage lending. *Review of Financial Studies*, **32** (5), 1854 – 1899.

GAMBACORTA, L., HUANG, Y., QIU, H. and WANG, J. (2019). How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. *Bank of Iinternational Settlements*, (mimeo).

GONZALEZ, L. and KOMOROVA LOUREIRO, Y. (2014). When can a photo increase credit? the impact of lender and borrower profiles on online peer-to-peer loans. *Journal of Behavioral and Experimental Finance*, **2**, 44–58.

HADDAD, C. and HORNUF, L. (2018). The emergence of the global fintech market: Economic and technological determinants. *Small business economics*, pp. 1– 25.

HAU, H., HUANG, Y., SHAN, H. and SHENG, Z. (2018). Fintech credit, financial inclusion, and enterpreneurial growth. *Working Paper*.

—, —, — and — (2019). How fintech enters China's credit market. *AEA Papers and Proceedings, American Economic Association*, **109**, 60–64.

HERZENSTEIN, M., SONENSHEIN, S. and DHOLAKIA, U. M. (2011). Tell me a good story and i may lend you money: The role of narratives in peer-to-peer lending decisions. *Journal of Marketing Research*, **48**, S138–S149.

IYER, R., KHWAJA, A. I., LUTTMER, E. F. P. and SHUE, K. (2016). Screening peers softly: Inferring the quality of small borrowers. *Management Science*, **62** (6), 1554–1577.

JAGTIANI, J. and LEMIEUX, C. (2018a). Do fintech lenders penetrate areas that are underserved by traditional banks? *Journal of Economics and Business*, (100), 43 – 54.

— and — (2018b). The roles of alternative data and machine learning in fintech lending: Evidence from the lendingclub consumer platform. *Working Papers, Federal Reserve Bank of Philadelphia*.

KHANDANI, A. E., KIM, A. J. and LO, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, **34** (11), 2767–2787.

LIN, M., PRABHALA, N. R. and VISWANATHAN, S. (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science*, **59** (1), 17–35.

MAGGIO, M. D. and YAO, V. (2019). Fintech borrowers: Lax-screening or cream-skimming? *SSRN Working Paper*, (3224957).

MCKINSEY (2019). The last pit stop? Time for bold late-cycle moves. *McKinsey Global Banking Annual Review*.

MORSE, A. (2015). Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending. *Annual Review of Financial Economics*, **7**, 463–482.

MOSCATELLI, M., NARIZZANO, S., PARLAPIANO, F. and VIGGIANO, G. (2019). Corporate default forecasting with machine learning. *forthcoming*.

NAVARETTI, G. B., CALZOLARI, G. and POZZOLO, A. F. (2017). Fintech and banks: Friends or foes? *European Economy: Banks, Regulation, and the Real Sector*, (December).

PETRALIA, K., PHILIPPON, T., RICE, T. and VERON, N. (2019). Banking disrupted? Financial intermediation in an era of transformational technology. *The Geneva Report on the World Economy*, (22).

POPE, D. G. and SYDNOR, J. R. (2011). What's in a picture? evidence of discrimination from prosper.com. *Journal of Human Resources*, **46** (1), 53–92.

PORTA, R. L., FLORENCIO LOPEZ-DE SILANES, A. S. and VISHNY, R. W. (1998). Law and finance. *Journal of Political Economy*, (6), 1113–1155.

POTZSCH, S. and BOHME, R. (2010). The role of soft information in trust building: Evidence from online social lending. *In: International Conference on Trust and Trustworthy Computing*, pp. 381–395.

RAU, R. (2018). Law, trust, and the development of crowdfunding. *SSRN Working Paper*, (2989056).

RAVINA, E. (2012). Love and loans: The effect of beauty and personal characteristics in credit markets? *SSRN Working Paper*.

TANG, H. (2019). Peer-to-peer lenders versus banks: Substitutes or complements? *The Review of Financial Studies*, **32** (5), 1900âĂŞ1938.

THAKOR, A. V. (2019). Fintech and banking: what do we know? *Journal of Financial Intermediation*, (In press).

WILSON, N., SUMMERS, B. and HOPE, R. (2000). Using payment behaviour data for credit risk modelling. *International journal of the Economics of Business*, **7** (3), 333–346.

WOLFE, B. and YOO, W. (2018). Crowding out banks: Credit substitution by peer-to-peer lending. *SSRN Working paper*, (3000593).